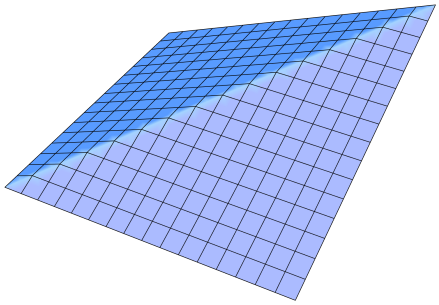


PMATH 334: Rings and Fields with Applications

Prof. Blake Madill • Winter 2016 • University of Waterloo
MWF 2:30–3:20PM in MC 2034

Transcribed, LaTeXed, and edited by Rob Zimmerman

Note: These are personal lecture notes, not official course materials. They may contain errors (which by default you should attribute to me, Rob Zimmerman, rather than the course instructor). The permanent home of these — and all of my other lecture notes — is https://rob-zimmerman.github.io/coursenotes.

Contents

Lecture 1 — Monday, January 04, 2016

1. Basic Ring Theory

Rings and Familiar Examples

We begin with a few familiar examples.

$\mathbb{Z}$	$\mathbb{C}$	$M_{2\times 2}(\mathbb{R})$	$\mathbb{Q}[x]$	$2\mathbb{Z}$	
$+,\cdot$	$+,\cdot$	$+,\cdot$	$+,\cdot$	$+,\cdot$	Operations
0	0	$\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$	0	0	Additive identity
1	1	$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$	1	none	Multiplicative identity

The goal of the course is to generalize familiar sets equipped with addition and multiplication. Along the way, this also generalizes vector spaces, polynomial factorization, prime factorizations, congruences and the Chinese remainder theorem, and linear transformations (among many others).

Definition 1.1 (Ring) A **ring** R is a nonempty set equipped with two operations: **addition** $+ : R \times R \to R$ and **multiplication** $\cdot : R \times R \to R$. We write $+(a, b)$ as $a + b$, and we write $\cdot(a, b)$ as $a \cdot b$ or ab. These operations satisfy the following **ring axioms**:

1. (**Commutativity of addition**) $a + b = b + a$ for all $a, b \in R$.
2. (**Associativity of addition**) $(a + b) + c = a + (b + c)$ for all $a, b, c \in R$.
3. (**Existence of additive identity**) There exists an element $0 \in R$ such that $0 + a = a + 0 = a$ for all $a \in R$.
4. (**Existence of additive inverse**) For each $a \in R$, there exists an element $-a \in R$ such that $a + (-a) = 0$.
5. (**Associativity of multiplication**) $a(bc) = (ab)c$ for all $a, b, c \in R$.
6. (**Distributivity**) $a(b + c) = ab + ac$ and $(b + c)a = ba + ca$ for all $a, b, c \in R$.

Notation 1.2 We call $+$ the **ring addition** and $\cdot$ the **ring multiplication**. We call 0 the **additive identity**, and we call $-a$ the **additive inverse** of $a \in R$.

In other words, a ring is an abelian group under addition together with a multiplication operation that distributes over addition.

Remark 1.3

1. We do not require ring multiplication to be commutative.
2. An element 1 such that $1a = a1 = a$ for all $a \in R$ need not exist, but if it does, we call it the **unity** of the ring.

Example 1.4 The sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $M_{n\times n}(\mathbb{R})$, $\mathbb{Q}[x]$, and $2\mathbb{Z}$ are all rings. More generally, if R is a ring, then

$$M_n(R) = \{n \times n \text{ matrices over } R\}$$

is a ring with matrix addition and multiplication defined using the operations of R. Explicitly, if $A = (a_{ij})$ and $B = (b_{ij})$ belong to $M_n(R)$, then

$$(A+B)_{ij} = a_{ij} + b_{ij} \qquad \text{and} \qquad (AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Example 1.5 The set $\mathbb{Z}[i]$ of Gaussian integers is a ring under the usual complex addition and multiplication. The set $2\mathbb{Z} = \{2k : k \in \mathbb{Z}\}$ is another example. Also,

$$C(\mathbb{R},\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : f \text{ is continuous}\}$$

is a ring under pointwise addition and multiplication:

$$(f+g)(x) = f(x) + g(x) \qquad \text{and} \qquad (f \cdot g)(x) = f(x)g(x).$$

Example 1.6 The following familiar-looking sets are not rings:

1. $\mathbb{N} = \{1, 2, \dots\}$ with the usual operations is not a ring because $0 \notin \mathbb{N}$.
2. $S = \{0, 1, 2, \dots\}$ with the usual operations is not a ring because its nonzero elements do not have additive inverses.
3. $S = \{2k+1 : k \in \mathbb{Z}\}$ with the usual operations is not a ring because it has no additive identity.
4. $S = \left\{ \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} : a, b \in \mathbb{C} \right\}$ with the usual matrix operations is not a ring because it is

not closed under multiplication. For example,

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \notin S.$$

Lecture 2 — Wednesday, January 06, 2016

Recall that a ring R is a nonempty set with an addition $+$ and a multiplication $\cdot$ satisfying the ring axioms.

Remark 2.1 These are *not* ring axioms:

1. $ab = ba$ for all $a, b \in R$.
2. There exists $1 \in R$ such that $1a = a1 = a$ for all $a \in R$.
3. For every $a \in R$, there exists $b \in R$ such that $ab = 1$.

Notation 2.2 Let $a, b \in R$. Instead of writing $a + (-b)$, we write $a - b$.

Special Kinds of Rings

Definition 2.3 Let R be a ring. If $ab = ba$ for all $a, b \in R$, then we say that R is a **commutative ring**. Otherwise, we say that R is a **noncommutative ring**. If there exists $1 \in R$ such that $a1 = 1a = a$ for all $a \in R$, then we say that R is a **unital ring**. We call the element 1 the **unity** of the ring.

Example 2.4 Let $a, b \in \mathbb{Z}$ and let $n \in \mathbb{Z}$ with $n \geq 1$. We say that $a \equiv b \pmod{n}$ if $n \mid (a - b)$. For example, $5 \equiv 2 \pmod{3}$. Let

$$[a]_n = \{b \in \mathbb{Z} : a \equiv b \pmod{n}\},$$

the congruence class of a modulo n. By the division algorithm, there exist $q, r \in \mathbb{Z}$ such that $a = nq + r$ and $r \in \{0, 1, \ldots, n-1\}$. Then $a \equiv r \pmod{n}$. Therefore, for every $a \in \mathbb{Z}$, there

exists b with $0 \leq b < n$ such that $a \equiv b \pmod{n}$. Let

$$\mathbb{Z}_n = \{[0], [1], \ldots, [n-1]\},$$

which we call the **integers modulo n**.

Define

$$[a] + [b] = [a+b], \qquad \text{and} \qquad [a] \cdot [b] = [ab].$$

These operations are well-defined. Indeed, suppose that $[a] = [a']$ and $[b] = [b']$. Then $n \mid (a - a')$ and $n \mid (b - b')$, so

$$n \mid (a+b) - (a'+b').$$

Also,

$$ab - a'b' = b(a - a') + a'(b - b'),$$

so $n \mid (ab - a'b')$. Hence $[a+b] = [a'+b']$ and $[ab] = [a'b']$.

These operations make $\mathbb{Z}_n$ into a ring. Moreover, $\mathbb{Z}_n$ is a **finite ring**, meaning that it has finitely many elements.

Notation 2.5 Instead of writing $[a]_n$, we simply write a. The value of n is understood from context.

Example 2.6 In $\mathbb{Z}_6$, we have $2 + 3 = 5$. In $\mathbb{Z}_5$, we have $2 + 4 = 1$.

These examples compare as follows:

Rings	0	Commutative	Unital	1
$\mathbb{Z}[i]$	$0 + 0i$	✓	✓	$1 + 0i$
$M_2(\mathbb{Z}_2)$	$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	×	✓	$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
$C(\mathbb{R}, \mathbb{R})$	$f \equiv 0$	✓	✓	$f \equiv 1$
$2\mathbb{Z}[x]$	0	✓	×	none

Example 2.7 Let $R = \{0\}$ and define $0 + 0 = 0$ and $0 \cdot 0 = 0$. This is a ring, called the **trivial ring**.

Remark 2.8 Let $(R, +, \cdot)$ be a ring. We can construct a new ring R' by taking $R' = R$ as a set, keeping the same addition, and defining $a \cdot b = 0$ for all $a, b \in R'$. We call R' a **ring with trivial multiplication**. Unless otherwise stated, rings have nontrivial multiplication. We do, however, consider the trivial ring to be a ring.

Proposition 2.9 Let R be a ring. Let $a, b, c \in R$. Then

1. $a0 = 0a = 0$.
2. $a(-b) = (-a)b = -(ab)$.
3. $(-a)(-b) = ab$.
4. $a(b - c) = ab - ac$.

Proof. For (1), since

$$a0 = a(0 + 0) = a0 + a0,$$

adding $-(a0)$ to both sides gives $0 = a0$. Similarly, $0a = 0$. For (2), we have

$$(-a)b + ab = (-a + a)b = 0b = 0,$$

so $(-a)b = -(ab)$. Similarly, $a(-b) = -(ab)$. For (3), by (2),

$$(-a)(-b) = -(a(-b)) = -(-(ab)).$$

Since the additive inverse is unique, $-(-(ab)) = ab$. For (4), we compute

$$a(b - c) = a(b + (-c)) = ab + a(-c) = ab - (ac) = ab - ac.$$

□

Proposition 2.10 The additive identity of a ring R is unique.

Proof. Suppose that 0 and $\tilde{0}$ are both additive identities. Since $\tilde{0}$ is an additive identity, $0 + \tilde{0} = 0$. Since 0 is an additive identity, $0 + \tilde{0} = \tilde{0}$. Hence $0 = \tilde{0}$. □

Proposition 2.11 Let R be a unital ring. Then the unity of R is unique.

Proof. Suppose that 1 and $\tilde{1}$ are both unities. Since $\tilde{1}$ is a unity, $1 \cdot \tilde{1} = 1$. Since 1 is a unity, $1 \cdot \tilde{1} = \tilde{1}$. Hence $1 = \tilde{1}$. □

Subrings

Definition 2.12 Let R be a ring. A **subring** of R is a nonempty subset of R which forms a ring under the operations of R.

Example 2.13 $\mathbb{Z}$ is a subring of both $\mathbb{C}$ and $\mathbb{R}$. $\{0\}$ is a subring of any ring. $\mathbb{Q}$ is a subring of $\mathbb{Q}[x]$.

Theorem 2.14 (Subring test) A nonempty subset S of a ring R is a subring of R if

1. $a - b \in S$ for all $a, b \in S$.
2. $ab \in S$ for all $a, b \in S$.

The subring test usually makes it much easier to prove that a subset is a ring, because it lets us use the fact that the ambient ring is already known to be a ring.

As Figure 1 emphasizes, almost all of the ring axioms are inherited from the ambient ring. After checking nonemptiness, only closure under subtraction and multiplication remains.

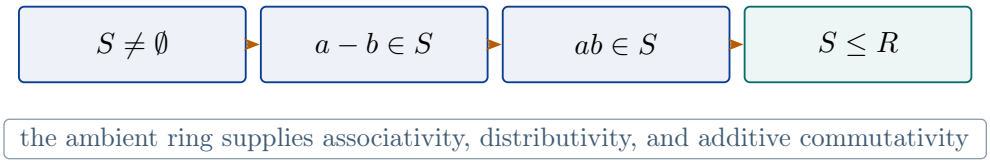


FIGURE 1

Lecture 3 — Friday, January 08, 2016

Proof of Theorem 2.14. Let $S \neq \emptyset$ and suppose that S satisfies (1) and (2). The commutativity of addition, associativity of addition, associativity of multiplication, and distributivity are inherited from R. For any $a \in S$, we have $a - a = 0 \in S$, so S contains the additive identity. If $a \in S$, then $0, a \in S$, and hence $0 - a = -a \in S$. Thus every element of S has its additive inverse in S.

It remains to check closure under addition and multiplication. Closure under multiplication is exactly (2). If $a, b \in S$, then $-b \in S$, so (1) gives

$$a + b = a - (-b) \in S.$$

Thus S is a ring under the operations of R. □

Example 3.1 We show that

$$T = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} : a, b, c \in \mathbb{R} \right\}$$

is a subring of $M_2(\mathbb{R})$. Since the zero matrix belongs to T, the set T is nonempty. For $a, b, c, \bar{a}, \bar{b}, \bar{c} \in \mathbb{R}$, we have

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} \bar{a} & \bar{b} \\ 0 & \bar{c} \end{pmatrix} = \begin{pmatrix} a\bar{a} & a\bar{b} + b\bar{c} \\ 0 & c\bar{c} \end{pmatrix} \in T$$

and

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} - \begin{pmatrix} \bar{a} & \bar{b} \\ 0 & \bar{c} \end{pmatrix} = \begin{pmatrix} a - \bar{a} & b - \bar{b} \\ 0 & c - \bar{c} \end{pmatrix} \in T.$$

By Theorem 2.14, T is a subring of $M_2(\mathbb{R})$.

Example 3.2 Let R be a ring and take $r \in R$. The **centralizer** of r is

$$C(r) = \{a \in R : ar = ra\}.$$

We show that $C(r)$ is a subring. Since $0r = 0 = r0$, we have $0 \in C(r)$, so $C(r)$ is nonempty.

If $a, b \in C(r)$, then

$$(a-b)r = ar - br = ra - rb = r(a-b),$$

so $a - b \in C(r)$. Also,

$$(ab)r = a(br) = a(rb) = (ar)b = (ra)b = r(ab),$$

so $ab \in C(r)$. By Theorem 2.14, $C(r)$ is a subring of R.

Warning Subrings of unital rings need not be unital. Textbooks differ on whether a subring of a unital ring is required to contain the same unity.

Example 3.3 The ring $2\mathbb{Z}$ is a subring of $\mathbb{Z}$. The ring $\mathbb{Z}$ is unital, but $2\mathbb{Z}$ is not.

Example 3.4 In $\mathbb{Z}_6 = \{0, 1, 2, \ldots, 5\}$, the subset $S = \{0, 2, 4\}$ is a subring. At first glance S may appear not to be unital, but 4 is a unity for S:

$$0 \cdot 4 = 0, \qquad 2 \cdot 4 = 2, \qquad \text{and} \qquad 4 \cdot 4 = 4.$$

Example 3.5 Consider $\mathbb{Z}$ with the usual addition and a new multiplication defined by $a \cdot b = -ab$. The subring test cannot be used to prove that this is a ring, because the multiplication is not inherited from the usual ring structure on $\mathbb{Z}$. Instead, we must check the ring axioms directly. This ring is unital, with unity -1.

Direct Sums

Definition 3.6 Let $R_1, R_2, \ldots, R_n$ be rings. The **direct sum** of $R_1, \ldots, R_n$ is

$$R_1 \oplus R_2 \oplus \cdots \oplus R_n = \{(a_1, a_2, \ldots, a_n) : a_i \in R_i \text{ for } 1 \leq i \leq n\},$$

with operations defined componentwise:

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$$

and

$$(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = (a_1 b_1, \ldots, a_n b_n).$$

The operations in each component are the operations of the corresponding ring R_i. With these operations, the direct sum is a ring.

The componentwise character of both operations is pictured in Figure 2. No coordinate interacts with a different coordinate; this is why properties such as commutativity can be checked one ring at a time.

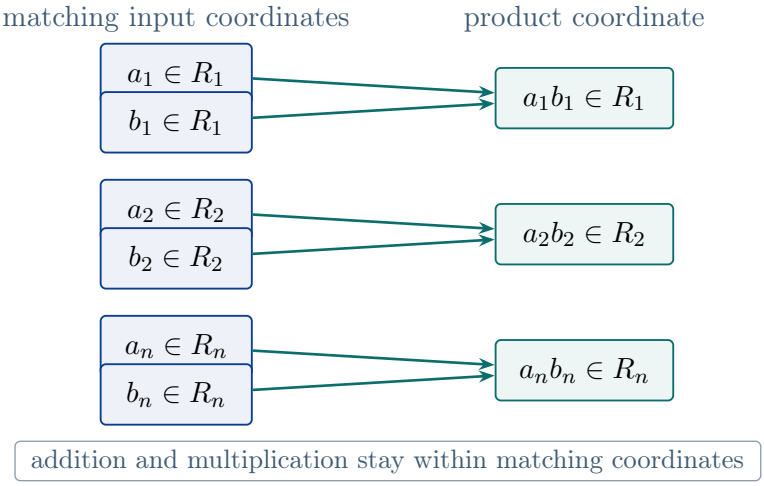


FIGURE 2

Example 3.7 In $\mathbb{C} \oplus \mathbb{Z}_4$, we have

$$(1 + i, 2) \cdot (4i, 3) = (4i - 4, 2).$$

Example 3.8

1. Prove that $R_1 \oplus \cdots \oplus R_n$ is commutative if and only if each R_i is commutative.
2. Prove that $R_1 \oplus \cdots \oplus R_n$ is unital if and only if each R_i is unital.

Solution. For (1), we prove both implications.

Forward implication. Suppose that $R_1 \oplus \cdots \oplus R_n$ is commutative. Fix i, and let $a, b \in R_i$. Let x be the element of the direct sum whose ith component is a and whose other components are 0,

and let y be the element whose ith component is b and whose other components are 0. Since the direct sum is commutative, $xy = yx$. Comparing the ith components gives $ab = ba$. Thus R_i is commutative.

Reverse implication. Suppose that each R_i is commutative. If

$$x = (a_1, \ldots, a_n) \qquad \text{and} \qquad y = (b_1, \ldots, b_n),$$

then

$$xy = (a_1 b_1, \ldots, a_n b_n) = (b_1 a_1, \ldots, b_n a_n) = yx.$$

Therefore $R_1 \oplus \cdots \oplus R_n$ is commutative.

For (2), we again prove both implications.

Forward implication. Suppose that $R_1 \oplus \cdots \oplus R_n$ has unity

$$e = (e_1, \ldots, e_n).$$

Fix i and let $a \in R_i$. Let x be the element of the direct sum whose ith component is a and whose other components are 0. Since $ex = xe = x$, comparing the ith components gives

$$e_i a = a e_i = a.$$

Thus e_i is a unity for R_i.

Reverse implication. Suppose that each R_i has unity 1_i. Then

$$(1_1, \ldots, 1_n)$$

is a unity for $R_1 \oplus \cdots \oplus R_n$, since multiplication is defined componentwise. □

Integral Domains and Fields

Rings generalize many familiar algebraic objects, including $\mathbb{C}$, $\mathbb{Z}$, and $\mathbb{Z}_n$. However, not every ring is a good analogue of $\mathbb{Z}$, $\mathbb{R}$, or $\mathbb{C}$. Integral domains give a better generalization of $\mathbb{Z}$, while fields generalize examples such as $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$.

Definition 3.9 A **zero divisor** of a ring R is a nonzero element $a \in R$ such that there exists a nonzero element $b \in R$ with $ab = 0$.

Definition 3.10 An **integral domain** is a nontrivial commutative unital ring with no zero divisors.

Example 3.11 The ring $\mathbb{Z}$ is an integral domain. The ring $\mathbb{Z}_4$ is not, since $2 \cdot 2 = 0$ in $\mathbb{Z}_4$ but $2 \neq 0$.

Example 3.12 The ring $M_2(\mathbb{R})$ is not commutative, so it is not an integral domain.

Example 3.13 The ring $\mathbb{Z}_6 \oplus \mathbb{Z}_6$ is not an integral domain, since

$$(0, 4) \cdot (2, 0) = (0, 0).$$

Lecture 4 — Monday, January 11, 2016

Recall that a nonzero element $a \in R$ is a zero divisor if there exists $0 \neq b \in R$ such that $ab = 0$. A ring R is an integral domain if it is commutative, unital, and has no zero divisors. Thus $\mathbb{Z}$, $\mathbb{C}$, $\mathbb{Q}$, and $\mathbb{R}$ are all integral domains.

Equivalently, a commutative unital ring is an integral domain if and only if $ab = 0$ implies that $a = 0$ or $b = 0$. Unital subrings of integral domains are themselves integral domains. For example, $\mathbb{Z}[i]$ is an integral domain because it is a unital subring of $\mathbb{C}$. On the other hand, $\mathbb{Z}_4$ is not an integral domain, since $2 \cdot 2 = 0$ in $\mathbb{Z}_4$.

Proposition 4.1 Let $n \in \mathbb{Z}$ with $n \geq 2$. Then a nonzero element $a \in \mathbb{Z}_n$ is a zero divisor if and only if $\gcd(a, n) \neq 1$.

Proof. *Forward implication.* Suppose that $a \in \mathbb{Z}_n \setminus \{0\}$ is a zero divisor. Then there exists $0 \neq b \in \mathbb{Z}_n$ such that $ab = 0$ in $\mathbb{Z}_n$. Hence $n \mid ab$ in $\mathbb{Z}$. If $\gcd(a, n) = 1$, then Euclid's lemma gives $n \mid b$, so $b = 0$ in $\mathbb{Z}_n$, a contradiction. Therefore $\gcd(a, n) \neq 1$.

Reverse implication. Suppose that $\gcd(a, n) = d \neq 1$. Then $d \mid n$ and $d \mid a$, so $n = dk$ and $a = dl$ for some $k, l \in \mathbb{Z}$. Since $d > 1$, we have $0 < k < n$, and hence $k \neq 0$ in $\mathbb{Z}_n$. In $\mathbb{Z}_n$,

$$ak = dlk = nl = 0.$$

Thus a is a zero divisor. □

Corollary 4.2 The ring $\mathbb{Z}_n$ is an integral domain if and only if n is prime.

Proof. If $\mathbb{Z}_n$ is an integral domain, then it has no zero divisors. By Proposition 4.1, every nonzero residue class in $\mathbb{Z}_n$ is relatively prime to n. Therefore no integer a with $1 < a < n$ can divide n, and so n is prime.

Conversely, if n is prime, then every nonzero $a \in \mathbb{Z}_n$ satisfies $\gcd(a, n) = 1$. By Proposition 4.1, $\mathbb{Z}_n$ has no zero divisors. Since $\mathbb{Z}_n$ is commutative and unital, it is an integral domain. □

Definition 4.3 Let R be a unital ring. We say that $a \in R$ is a **unit**, or **invertible**, if there exists $b \in R$ such that $ab = ba = 1$. We denote b by a^{-1}.

Definition 4.4 A **field** is a nontrivial commutative unital ring such that every nonzero element is invertible, that is, every nonzero element is a unit.

Example 4.5 The rings $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Q}$ are all fields. The ring $\mathbb{Z}$ is not a field, because 2 is not invertible in $\mathbb{Z}$.

Example 4.6 For $n > 1$, the ring $M_n(\mathbb{R})$ is not a field. Indeed, a matrix $A \in M_n(\mathbb{R})$ is a unit if and only if $\det(A) \neq 0$. Moreover, $M_n(\mathbb{R})$ is noncommutative.

Example 4.7 Consider

$$\mathbb{Z}_3[i] = \{a + bi : a, b \in \mathbb{Z}_3,\ i^2 = -1\}.$$

For instance,

$$(1 + 2i)(1 + i) = 1 + i + 2i - 2 = -1 + 3i = 2.$$

The ring $\mathbb{Z}_3[i]$ is a field of size $3 \cdot 3 = 9$. Its addition and multiplication are commutative, and 1 is a unity. Its nonzero elements have inverses as follows:

$$1 \cdot 1 = 1, \qquad 2 \cdot 2 = 1, \qquad i(2i) = 1, \qquad \text{and} \qquad 2i \cdot i = 1,$$

$(1+i)(2+i) = 1, \qquad (1+2i)(2+2i) = 1, \qquad (2+i)(1+i) = 1, \qquad \text{and} \qquad (2+2i)(1+2i) = 1.$

Notation 4.8 We denote the set of units of a ring R by $\mathcal{U}(R)$. The set $\mathcal{U}(R)$ is a group under ring multiplication, called the **group of units**.

Example 4.9 Let $R_1, \ldots, R_n$ be commutative and unital rings. Then

$$\mathcal{U}(R_1 \oplus \cdots \oplus R_n) = \{(a_1, \ldots, a_n) : a_i \in \mathcal{U}(R_i) \text{ for each } i\}.$$

Definition 4.10 A ring R has **right cancellation** if $ab = cb$ implies $a = c$ for all $a, b, c \in R$ with $b \neq 0$. Similarly, R has **left cancellation** if $ba = bc$ implies $a = c$ for all $a, b, c \in R$ with $b \neq 0$. We say that R is **cancellative** if it has both left and right cancellation.

Example 4.11 The ring $\mathbb{Z}_6$ does not have right cancellation, since $2 \cdot 3 = 0 \cdot 3$ but $2 \neq 0$.

Proposition 4.12 Integral domains are cancellative.

Proof. Let R be an integral domain. Suppose that $ab = ac$ with $a \neq 0$. Then $ab - ac = 0$, so $a(b - c) = 0$. Since R is an integral domain and $a \neq 0$, we have $b - c = 0$, and hence $b = c$. Therefore R has left cancellation. Since R is commutative, it also has right cancellation. Thus R is cancellative. □

Proposition 4.13 Every field is an integral domain (and thus cancellative).

Proof. Let K be a field. Then K is a commutative unital ring by definition, so it remains to show that K has no zero divisors.

Suppose that $ab = 0$ for some $a, b \in K$. If $b = 0$, there is nothing to prove. Otherwise, b has an inverse b^{-1}. Then

$$abb^{-1} = 0b^{-1},$$

so $a \cdot 1 = 0$, and hence $a = 0$. Therefore K is an integral domain. □

Example 4.14 $\mathbb{Z}$ is an integral domain, but not a field.

Lecture 5 — Wednesday, January 13, 2016

2. Ideals, Factor Rings, and Homomorphisms

Recall that integral domains are commutative, fields are integral domains, and $\mathbb{Z}$ is not a field. Also, for $a \in \mathbb{Z}_n \setminus \{0\}$, Proposition 4.1 says that a is a zero divisor if and only if $\gcd(a, n) \neq 1$, and Corollary 4.2 says that $\mathbb{Z}_n$ is an integral domain if and only if n is prime.

Theorem 5.1 Every finite integral domain is a field.

Proof. Let R be a finite integral domain. Take $0 \neq a \in R$. We show that a^{-1} exists. If $a = 1$, then $a^{-1} = 1$, so suppose that $a \neq 1$. Consider the powers

$$a, a^2, a^3, a^4, \ldots.$$

Since R is finite, there exist integers $i > j$ such that $a^i = a^j$. By cancellativity, $a^{i-j} = 1$. Since $a \neq 1$, we have $i - j > 1$. Thus $aa^{i-j-1} = a^{i-j} = 1$, so $a^{-1} = a^{i-j-1}$. Therefore R is a field. □

Corollary 5.2 The ring $\mathbb{Z}_n$ is a field if and only if n is prime.

Proof. *Forward implication.* Suppose that $\mathbb{Z}_n$ is a field. Then $\mathbb{Z}_n$ is an integral domain by Proposition 4.13, so n is prime by Corollary 4.2.

Reverse implication. Suppose that n is prime. Then $\mathbb{Z}_n$ is a finite integral domain by Corollary 4.2, so $\mathbb{Z}_n$ is a field by Theorem 5.1. □

Proposition 5.3 Let $0 \neq a \in \mathbb{Z}_n$. Then a is a unit if and only if $\gcd(a, n) = 1$.

Proof. *Forward implication.* Suppose that a is a unit. Then a is not a zero divisor. By Proposi-

tion 4.1, we must have $\gcd(a, n) = 1$.

Reverse implication. Suppose that $\gcd(a, n) = 1$. By Bezout's identity, there exist $x, y \in \mathbb{Z}$ such that $ax + ny = 1$. In $\mathbb{Z}_n$, this gives $ax + 0 = 1$, so $ax = 1$. Hence a is invertible. □

Example 5.4 Every nonzero element of $\mathbb{Z}_n$ is either a unit or a zero divisor.

Indeed, if $0 \neq a \in \mathbb{Z}_n$, then either $\gcd(a, n) = 1$ or $\gcd(a, n) \neq 1$. In the first case, Proposition 5.3 shows that a is a unit. In the second case, Proposition 4.1 shows that a is a zero divisor.

Definition 5.5 Let R be a ring. The **characteristic** of R, denoted $\text{char}(R)$, is the least positive integer n such that $nr = 0$ for all $r \in R$. If no such n exists, then we say that $\text{char}(R) = 0$.

Example 5.6 We have

$$\text{char}(\mathbb{Z}) = 0, \qquad \text{char}(\mathbb{R}) = 0, \qquad \text{and} \qquad \text{char}(\mathbb{Z}_n) = n,$$

and also

$$\text{char}(\mathbb{Z}_n[x]) = n \qquad \text{and} \qquad \text{char}(M_n(\mathbb{Z}_n)) = n.$$

Let R be a commutative ring of characteristic p, where p is prime. Then $p \mid \binom{p}{k}$ for $k = 1, 2, \ldots, p-1$. Hence, for $a, b \in R$,

$$(a + b)^p = \sum_{i=0}^{p} \binom{p}{i} a^i b^{p-i} = a^p + b^p.$$

This identity is called the **freshman's dream**.

Proposition 5.7 Let R be an integral domain. Then $\text{char}(R)$ is either 0 or prime.

Proof. If $\text{char}(R) = 0$, there is nothing to prove. Suppose instead that $\text{char}(R) > 0$, say $\text{char}(R) = n$. For a contradiction, suppose that n is composite, say $n = kl$ for some $1 < k, l < n$. Then

$$0 = n \cdot 1 = (kl) \cdot 1 = (k \cdot 1)(l \cdot 1).$$

Since R is an integral domain, either $k \cdot 1 = 0$ or $l \cdot 1 = 0$. If $k \cdot 1 = 0$, then $kr = 0$ for all $r \in R$,

contradicting the minimality of n. The same contradiction follows if $l \cdot 1 = 0$. Therefore $\text{char}(R)$ is zero or prime. □

Example 5.8 We have $\text{char}(\mathbb{Z}_p) = p$ and $\text{char}(\mathbb{Z}) = 0$.

Proposition 5.9 If $\text{char}(R) = 0$, then R is infinite.

Proof. Suppose that R is finite. The additive multiples of each $r \in R$ must eventually repeat, so some positive integer n_r satisfies $n_r r = 0$. Because R is finite, we may choose a positive integer N divisible by every n_r. Then $Nr = 0$ for all $r \in R$, contradicting $\text{char}(R) = 0$. Therefore R is infinite. □

Ideals and Principal Ideals

Our next goal is to construct factor rings.

Definition 5.10 Let R be a ring. A subring S of R is a **left ideal** if $Ra \subseteq S$ for all $a \in S$. We say that S is a **right ideal** if $aR \subseteq S$ for all $a \in S$. If S is both a left ideal and a right ideal, then we call it an **ideal**.

Example 5.11 $2\mathbb{Z}$ is an ideal of $\mathbb{Z}$. Take $2k \in 2\mathbb{Z}$. Then for any $n \in \mathbb{Z}$, $2kn = 2(kn) \in 2\mathbb{Z}$.

Example 5.12 In any ring R, $\{0\}$ and R itself are both ideals.

Example 5.13 The ring $\mathbb{Z}$ is not an ideal of $\mathbb{Z}[x]$. Indeed, $2 \in \mathbb{Z}$ and $x \in \mathbb{Z}[x]$, but $2x \notin \mathbb{Z}$.

Proposition 5.14 (Ideal test) Let R be a ring and let $A \subseteq R$. Then A is an ideal of R if

1. $A \neq \emptyset$.
2. $a - b \in A$ for all $a, b \in A$.
3. $ar, ra \in A$ for all $a \in A$ and $r \in R$.

Proof. By (1), (2), and Theorem 2.14, the set A is a subring of R, since (3) with $r \in A$ in particular gives closure under multiplication. The same condition (3) says exactly that A absorbs multiplication by arbitrary elements of R on both sides. Therefore A is an ideal of R. □

Example 5.15 Let R be a commutative ring. The set of **nilpotent elements** of R,

$$\mathrm{Nil}(R) = \{a \in R : a^k = 0 \text{ for some } k \in \mathbb{N}\}$$

is an ideal of R. Certainly $0 \in \mathrm{Nil}(R)$. If $a^m = 0$ and $b^n = 0$, then every term in the binomial expansion of

$$(a-b)^{m+n-1}$$

contains either a factor a^m or a factor b^n. Hence $(a-b)^{m+n-1} = 0$, so $a - b \in \mathrm{Nil}(R)$.

Now let $a \in \mathrm{Nil}(R)$ and let $r \in R$. Since a is nilpotent, there exists $k \in \mathbb{N}$ such that $a^k = 0$. Because R is commutative,

$$(ar)^k = (ra)^k = r^k a^k = r^k \cdot 0 = 0.$$

Therefore $ar, ra \in \mathrm{Nil}(R)$. By Proposition 5.14, $\mathrm{Nil}(R)$ is an ideal of R.

Example 5.16 Let R be the ring of lower triangular 2×2 matrices over $\mathbb{C}$, and let

$$A = \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} : a \in \mathbb{C} \right\} \subseteq R.$$

The zero matrix belongs to A, so $A \neq \emptyset$.

For $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} b & 0 \\ 0 & 0 \end{pmatrix} \in A$, we have

$$\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} b & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a-b & 0 \\ 0 & 0 \end{pmatrix} \in A$$

and

$$\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} ab & 0 \\ 0 & 0 \end{pmatrix} \in A.$$

By Theorem 2.14, A is a subring of R.

Now take $\begin{pmatrix} x & 0 \\ z & w \end{pmatrix} \in R$. Then

$$\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x & 0 \\ z & w \end{pmatrix} = \begin{pmatrix} ax & 0 \\ 0 & 0 \end{pmatrix} \in A,$$

so A is a right ideal of R. However,

$$\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} \notin A.$$

Thus A is not a left ideal.

Lecture 6 — Friday, January 15, 2016

Recall that a subring I of a ring R is an ideal if $ar, ra \in I$ for all $a \in I$ and $r \in R$. For example, $i\mathbb{R}$ is not an ideal of $\mathbb{C}$, since $i \in i\mathbb{R}$ and $i \in \mathbb{C}$, but

$$i \cdot i = -1 \notin i\mathbb{R}.$$

If I is an ideal of R, then we write $I \trianglelefteq R$, in analogy with the notation for normal subgroups.

Let R be a commutative unital ring. For $a_1, \ldots, a_n \in R$, define

$$(a_1, \ldots, a_n) = \{a_1 r_1 + a_2 r_2 + \cdots + a_n r_n : r_i \in R\}.$$

In the context of ideals, this is *not* a tuple! We also write $\langle a_1, \ldots, a_n \rangle = (a_1, \ldots, a_n)$.

Proposition 6.1 Let R be a commutative unital ring and let $a_1, \ldots, a_n \in R$. Then $(a_1, \ldots, a_n)$ is an ideal of R.

Proof. The set $(a_1, \ldots, a_n)$ is nonempty because $0 = a_1 \cdot 0 + \cdots + a_n \cdot 0$ belongs to it. If

$$x = a_1 r_1 + \cdots + a_n r_n \qquad \text{and} \qquad y = a_1 s_1 + \cdots + a_n s_n$$

belong to $(a_1, \ldots, a_n)$, then

$$x - y = a_1(r_1 - s_1) + \cdots + a_n(r_n - s_n) \in (a_1, \ldots, a_n).$$

If $r \in R$, then, using commutativity,

$$rx = a_1(rr_1) + \cdots + a_n(rr_n) \in (a_1, \ldots, a_n),$$

and similarly $xr \in (a_1, \ldots, a_n)$. Therefore $(a_1, \ldots, a_n)$ is an ideal by Proposition 5.14. □

We call this the **ideal generated by** $a_1, \ldots, a_n$. In particular, $a_i \in (a_1, \ldots, a_n)$ for all $1 \leq i \leq n$. In the context of ideals, $(a_1, \ldots, a_n)$ denotes an ideal, not a tuple.

Proposition 6.2 Let R be a commutative unital ring and let $a_1, \ldots, a_n \in R$. Then $(a_1, \ldots, a_n)$ is the smallest ideal of R containing each a_i.

Proof. The ideal $(a_1, \ldots, a_n)$ contains every a_i, because the ring is unital. Now let J be any ideal containing $a_1, \ldots, a_n$, and take

$$a_1r_1 + \cdots + a_nr_n \in (a_1, \ldots, a_n).$$

Since $a_i \in J$ and J is an ideal of R, each a_ir_i belongs to J. Hence $a_1r_1 + \cdots + a_nr_n \in J$, so $(a_1, \ldots, a_n) \subseteq J$. Therefore $(a_1, \ldots, a_n)$ is contained in every ideal that contains all the a_i, and is consequently the smallest such ideal. □

Remark 6.3 For a single element $a \in R$, we have

$$(a) = \{ar : r \in R\}.$$

For example, in $\mathbb{Z}$,

$$(n) = \{nz : z \in \mathbb{Z}\} = n\mathbb{Z}.$$

In $\mathbb{Z}[x]$,

$$(x) = \{x \cdot f(x) : f(x) \in \mathbb{Z}[x]\} = \{f(x) \in \mathbb{Z}[x] : f(0) = 0\},$$

the set of polynomials with no constant term.

Question 6.4 What are the ideals of $\mathbb{Z}$?

Recall the division algorithm: if $a, b \in \mathbb{Z}$ and $b \neq 0$, then there exist unique $q, r \in \mathbb{Z}$ such that $0 \leq r < |b|$ and $a = qb + r$.

Theorem 6.5 Every ideal of $\mathbb{Z}$ is of the form (a) for some $a \in \mathbb{Z}$.

Proof. Let $I \trianglelefteq \mathbb{Z}$. If $I = \{0\}$, then $I = (0)$. Suppose that $I \neq \{0\}$. Since I contains a nonzero integer and is closed under additive inverses, it contains a positive integer. Let a be the smallest positive integer in I. If $b \in I$, then by the division algorithm there exist $q, r \in \mathbb{Z}$ such that $b = aq + r$ and $0 \leq r < a$. Since $r = b - aq \in I$, the minimality of a forces $r = 0$. Hence $b = aq \in (a)$, so $I \subseteq (a)$. The reverse inclusion is clear because $a \in I$ and I is an ideal. Therefore $I = (a)$. □

Definition 6.6 An ideal of a commutative unital ring R is called a **principal ideal** if it is of the form (a) for some $a \in R$.

Theorem 6.5 says exactly that every ideal of $\mathbb{Z}$ is a principal ideal.

Example 6.7 Show that every ideal of $\mathbb{Z}_n$ is principal.

Solution. Let $I \trianglelefteq \mathbb{Z}_n$, and let $\pi : \mathbb{Z} \to \mathbb{Z}_n$ be the canonical map $\pi(a) = [a]_n$. Define

$$J = \{a \in \mathbb{Z} : [a]_n \in I\}.$$

Then J is an ideal of $\mathbb{Z}$. Indeed, J is nonempty because $0 \in J$. If $a, b \in J$, then $[a]_n, [b]_n \in I$, so

$$[a - b]_n = [a]_n - [b]_n \in I,$$

and hence $a - b \in J$. If $a \in J$ and $r \in \mathbb{Z}$, then

$$[ra]_n = [r]_n[a]_n \in I,$$

so $ra \in J$.

By Theorem 6.5, there is a nonnegative integer d such that $J = (d)$. Since $n \in J$, we have $d \mid n$ when $d \neq 0$. We claim that

$$I = ([d]_n).$$

If $[a]_n \in I$, then $a \in J = (d)$, so $a = dq$ for some $q \in \mathbb{Z}$. Hence

$$[a]_n = [d]_n[q]_n \in ([d]_n).$$

Conversely, since $d \in J$, we have $[d]_n \in I$, and therefore every multiple of $[d]_n$ lies in I. Thus $I = ([d]_n)$, so every ideal of $\mathbb{Z}_n$ is principal. □

Proposition 6.8 If $(a) \trianglelefteq \mathbb{Z}_n$, then there exists $b \in \mathbb{Z}_n$ such that $b \mid n$ in $\mathbb{Z}$ and $(a) = (b)$.

Proof. Let $b = \gcd(a, n)$, computed in $\mathbb{Z}$. Then $b \mid n$. Since $b \mid a$, we have $a = bc$ for some $c \in \mathbb{Z}$, so $a \in (b)$ and hence $(a) \subseteq (b)$ in $\mathbb{Z}_n$. Conversely, Bezout's identity gives $x, y \in \mathbb{Z}$ such that $b = ax + ny$. In $\mathbb{Z}_n$, this says $b = ax$, so $b \in (a)$ and hence $(b) \subseteq (a)$. Therefore $(a) = (b)$. □

Example 6.9 The ideals of $\mathbb{Z}_{12}$ are generated by the divisors of 12, with their inclusion lattice shown in Figure 3. In particular,

$$(1) = \{1 \cdot r : r \in \mathbb{Z}_{12}\} = \mathbb{Z}_{12}.$$

The ideals (2) and (3) are the prime and maximal ideals of $\mathbb{Z}_{12}$.

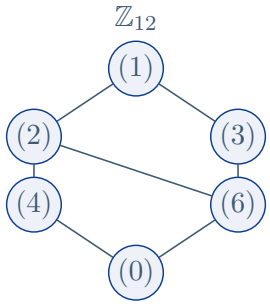


FIGURE 3

Example 6.10 Take $R = \mathbb{Z}[x]$ and let

$$J = (x^2) = \{a_2x^2 + a_3x^3 + \cdots + a_nx^n : n \geq 2,\ a_i \in \mathbb{Z}\}.$$

Also let

$$I = \{a_0 + a_1x + \cdots + a_nx^n : a_0 = a_1 = a_3 = 0\}.$$

Then $J \trianglelefteq R$. We claim that $I \trianglelefteq J$. Addition and subtraction preserve the conditions on the

coefficients defining I, while the product of two elements of I has degree at least 4 unless it is zero. Thus I is a subring of J. If

$$a_2x^2 + \cdots + a_nx^n \in I \qquad \text{and} \qquad b_2x^2 + b_3x^3 + \cdots + b_mx^m \in J,$$

then

$$(a_2x^2 + \cdots + a_nx^n)(b_2x^2 + \cdots + b_mx^m) = a_2b_2x^4 + \text{terms of higher order} \in I.$$

Thus $I \trianglelefteq J$. However,

$$(x^2 + x^4)(1 + x) = x^2 + x^3 + x^4 + x^5 \notin I,$$

so $I \not\trianglelefteq R$. Therefore being an ideal is not transitive.

Lecture 7 — Monday, January 18, 2016

Prime and Maximal Ideals

Proposition 7.1 Every ring with left cancellation also has right cancellation.

Proof. Suppose that R has left cancellation. Let $ab = cb$ with $b \neq 0$. Then

$$(a - c)b = 0 = (a - c)0.$$

If $a - c \neq 0$, then left cancellation gives $b = 0$, a contradiction. Therefore $a - c = 0$, so $a = c$. □

Remark 7.2 A ring R can have an element that is left cancellative but not right cancellative.

Recall the ideal lattice of $\mathbb{Z}_{12}$ shown in Figure 4.

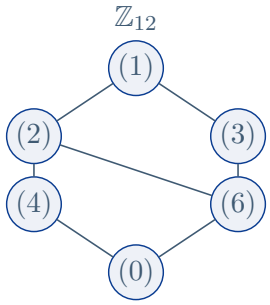


FIGURE 4

Definition 7.3 Let R be a commutative unital ring. A proper ideal P of R is **prime** if whenever $a, b \in R$ and $ab \in P$, we have $a \in P$ or $b \in P$.

Example 7.4 If p is prime, then $p\mathbb{Z} = (p)$ is a prime ideal of $\mathbb{Z}$. Indeed, suppose that $ab \in (p)$. Then $ab = pn$ for some $n \in \mathbb{Z}$, so $p \mid ab$. By Euclid's lemma, $p \mid a$ or $p \mid b$, and hence $a \in (p)$ or $b \in (p)$.

Definition 7.5 Let R be commutative and unital. A proper ideal M of R is **maximal** if whenever $I \trianglelefteq R$ and

$$M \subseteq I \subseteq R,$$

we have $I = M$ or $I = R$.

Example 7.6 The ideal $p\mathbb{Z} = (p) \trianglelefteq \mathbb{Z}$ is also maximal. Suppose that $I \trianglelefteq \mathbb{Z}$ and

$$(p) \subseteq I \subseteq \mathbb{Z}.$$

If $(p) \neq I$, then there exists $r \in I$ with $r \notin (p)$, so $p \nmid r$. Since p is prime, $\gcd(p, r) = 1$. Therefore there exist $x, y \in \mathbb{Z}$ such that

$$\underbrace{px}_{\in (p) \subseteq I} + \underbrace{ry}_{\in I} = 1.$$

Thus $1 \in I$, and so $I = \mathbb{Z}$. Therefore (p) is maximal.

In general, to show that M is maximal, we assume that $I \trianglelefteq R$ satisfies $M \subseteq I \subseteq R$ and $M \neq I$, and then prove that $I = R$.

Proposition 7.7 Every maximal ideal is prime.

The quotient-ring viewpoint below will give the proof.

Example 7.8 Let R be commutative and unital.

1. The ideal (0) is prime if and only if R is an integral domain. Indeed, $ab \in (0)$ means $ab = 0$, so the definition of prime ideal becomes exactly the condition that $ab = 0$ implies $a = 0$ or $b = 0$.

2. The ideal $(0) \trianglelefteq R$ is maximal if and only if R is a field. If (0) is maximal and $0 \neq a \in R$, then
$$(0) \subsetneq (a) \subseteq R,$$
so $(a) = R$. Hence there exists $r \in R$ such that $ar = 1$, and a is a unit. Conversely, if R is a field, then its only ideals are (0) and R, so (0) is maximal.

3. The ideal $(x^2 + 2x + 1) \trianglelefteq \mathbb{R}[x]$ is not prime. Indeed,
$$(x + 1)(x + 1) = x^2 + 2x + 1 \in (x^2 + 2x + 1),$$
but $x + 1 \notin (x^2 + 2x + 1)$, since every nonzero polynomial in $(x^2 + 2x + 1)$ has degree at least 2.

4. The ideal $(x^2 + 2) \trianglelefteq \mathbb{Z}_3[x]$ is not prime. Since $1^2 + 2 = 3 = 0$ in $\mathbb{Z}_3$, we have $(x - 1) \mid x^2 + 2$ in $\mathbb{Z}_3[x]$. Dividing gives
$$\begin{array}{r} x + 1 \\ x - 1 \,\big)\, x^2 + 0x + 2 \\ \underline{x^2 + 2x} \quad \downarrow \\ x + 2 \\ \underline{x + 2} \\ 0 \end{array}$$
Therefore $x^2 + 2 = (x - 1)(x + 1)$. Since $\mathbb{Z}_3$ is a field, it is an integral domain, so $\deg(f(x)(x^2 + 2)) \geq 2$ for every nonzero $f(x) \in \mathbb{Z}_3[x]$. Thus $x - 1, x + 1 \notin (x^2 + 2)$.

5. Let $R = \mathbb{Z} \oplus \mathbb{Z}$ and
$$I = \{(0, a) : a \in \mathbb{Z}\}.$$

Then $I \trianglelefteq R$. If $(a,b),(a',b') \in R$ and

$$(a,b)(a',b') = (aa',bb') \in I,$$

then $aa' = 0$. Since $\mathbb{Z}$ is an integral domain, $a = 0$ or $a' = 0$. Hence $(a,b) \in I$ or $(a',b') \in I$, so I is prime. However, I is not maximal. Indeed,

$$J = \{(2a,b) : a,b \in \mathbb{Z}\}$$

is an ideal of R and $I \subsetneq J \subsetneq R$.

Theorem 7.9 (Division algorithm for $K[x]$) Let K be a field. For all $f(x), g(x) \in K[x]$ with $g(x) \neq 0$, there exist unique $q(x), r(x) \in K[x]$ such that

$$f(x) = g(x)q(x) + r(x),$$

where $r(x) = 0$ or $\deg(r) < \deg(g)$. Here $q(x)$ is called the **quotient polynomial** and $r(x)$ is called the **remainder polynomial**.

Proof. This is the usual polynomial long division algorithm over a field. At each step, the leading coefficient of $g(x)$ can be inverted in K, so the leading term of the current dividend can be cancelled. The process terminates because the degree strictly decreases at each step, and the final remainder has degree smaller than $\deg(g)$ or is zero. Uniqueness follows because if

$$gq + r = gq' + r'$$

with both r and r' equal to zero or of degree less than $\deg(g)$, then $g(q - q') = r' - r$. If $q \neq q'$, the left-hand side has degree at least $\deg(g)$, while the right-hand side has degree less than $\deg(g)$, a contradiction. Hence $q = q'$ and then $r = r'$. □

Example 7.10 Let $R = \mathbb{R}[x]$ and $I = (x+1)$. We claim that I is maximal. Suppose that $J \trianglelefteq R$ and

$$I \subsetneq J \subseteq R.$$

We show that $J = R$. Choose $f(x) \in J \setminus I$. By Theorem 7.9,

$$f(x) = (x+1)q(x) + r(x),$$

where $r(x) = 0$ or $\deg(r) < 1$. Since $f(x) \notin I$, the remainder is nonzero. Hence $r(x) \in \mathbb{R}\backslash\{0\}$, so $r(x)$ is a unit. Moreover,

$$r(x) = \underbrace{f(x)}_{\in J} - \underbrace{(x+1)q(x)}_{\in J} \in J.$$

Thus J contains a unit, so $J = R$. Hence I is maximal in $\mathbb{R}[x]$.

Theorem 7.11 Let K be a field. Every ideal of $K[x]$ is principal.

Proof. Let $I \trianglelefteq K[x]$. If $I = \{0\}$, then $I = (0)$, so suppose that $I \neq (0)$. Choose $f(x) \in I \setminus \{0\}$ of minimal degree.

Take $g(x) \in I$. By Theorem 7.9,

$$g(x) = f(x)q(x) + r(x)$$

for some $q(x), r(x) \in K[x]$, where $r(x) = 0$ or $\deg(r) < \deg(f)$. If $r(x) \neq 0$, then

$$r(x) = \underbrace{g(x)}_{\in I} - \underbrace{f(x)q(x)}_{\in I} \in I,$$

contradicting the minimality of $\deg(f)$. Hence $r(x) = 0$.

Thus $g(x) = f(x)q(x) \in (f(x))$, so $I \subseteq (f(x))$. Since $f(x) \in I$, we also have $(f(x)) \subseteq I$. Therefore $I = (f(x))$. □

Lecture 8 — Wednesday, January 20, 2016

Set Theory and Zorn's Lemma

We now prove that every proper ideal $I \subsetneq R$ in a commutative unital ring is contained in a maximal ideal.

Definition 8.1 Let X be a set and let $\leq$ be a relation on X. We say that $(X, \leq)$ is a **poset**, or **partially ordered set**, if $\leq$ satisfies the following properties:

1. (**Reflexivity**) $a \leq a$ for all $a \in X$.
2. (**Transitivity**) If $a \leq b$ and $b \leq c$, then $a \leq c$.
3. (**Antisymmetry**) If $a \leq b$ and $b \leq a$, then $a = b$.

Definition 8.2 We say that a subset S of a poset X is **totally ordered** if, for all $a, b \in S$, either $a \leq b$ or $b \leq a$.

Example 8.3 The poset $(\mathbb{N}, \leq)$ is totally ordered. On the other hand, let $S = \mathcal{P}(\{1, 2, 3\})$ with the relation $\subseteq$. Since

$$\{1, 2\} \not\subseteq \{1, 3\} \qquad \text{and} \qquad \{1, 3\} \not\subseteq \{1, 2\},$$

the poset $(S, \subseteq)$ is not totally ordered.

Definition 8.4 Let $(S, \leq)$ be a poset. A **maximal element** in S is an element $a \in S$ such that, for all $b \in S$, if $a \leq b$, then $a = b$.

Example 8.5 Let

$$S = \{\text{proper subsets of } \{1, 2, 3\}\}$$

and order S by $\subseteq$. The subset $\{1, 3\}$ is a maximal element. However, it is not a greatest element, because $\{2\} \not\subseteq \{1, 3\}$. Maximality only rules out a strictly larger comparable element; it does not require every other element to lie below it.

Theorem 8.6 (Zorn's lemma) Let $(X, \leq)$ be a nonempty poset. If every totally ordered subset of X has an upper bound, then X itself has a maximal element.

Theorem 8.7 Let R be a commutative unital ring. Then every proper ideal $I \subsetneq R$ is contained in a maximal ideal.

Proof. Let

$$S = \{J \trianglelefteq R : I \subseteq J \subsetneq R\}.$$

Since $I \in S$, the set S is nonempty. Order S by inclusion. Let $\{P_\alpha\}_{\alpha \in X}$ be a totally ordered subset of S, and set

$$A = \bigcup_{\alpha \in X} P_\alpha.$$

We first check that A is an ideal of R. It is nonempty because it contains I. If $a, b \in A$, then $a \in P_\alpha$ and $b \in P_\beta$ for some $\alpha, \beta \in X$. Since the family is totally ordered, either $P_\alpha \subseteq P_\beta$ or $P_\beta \subseteq P_\alpha$, so both a and b lie in one of these ideals. Hence $a - b \in A$. If $r \in R$ and $a \in A$, then $a \in P_\alpha$ for some α, so $ra, ar \in P_\alpha \subseteq A$. Thus $A \trianglelefteq R$.

For each α, we have $P_\alpha \subseteq A$. Suppose for a contradiction that $A = R$. Then $1 \in A$, so $1 \in P_\beta$ for some $\beta \in X$. This forces $P_\beta = R$, contradicting $P_\beta \in S$. Therefore $A \neq R$. Since $I \subseteq P_\alpha$ for every α, we also have $I \subseteq A$, so $A \in S$. Thus A is an upper bound for $\{P_\alpha\}_{\alpha \in X}$ in S.

By Theorem 8.6, the poset S has a maximal element; call it M. Then M is an ideal of R containing I. If $M \subseteq J \subsetneq R$ for some ideal J of R, then $J \in S$, and maximality of M in S gives $J = M$. Therefore M is a maximal ideal of R containing I. □

Remark 8.8 This is the usual pattern for applying Zorn's lemma, including cases where the order relation is $\supseteq$ rather than $\subseteq$.

Factor Rings and Cosets

Definition 8.9 Let R be a ring and let S be a subset of R. The **coset** of S in R containing a is

$$a + S = \{a + s : s \in S\}.$$

Proposition 8.10 Let R be a ring and let S be a subring of R. Then the following hold:

1. $0 \in a + S$ if and only if $a \in S$.
2. $a + S = S$ if and only if $a \in S$.
3. $a + S = b + S$ if and only if $a - b \in S$.

4. Either $a + S = b + S$ or $(a + S) \cap (b + S) = \emptyset$.

Proof. For (1), if $0 \in a + S$, then $0 = a + s$ for some $s \in S$, so $a = -s \in S$. Conversely, if $a \in S$, then $-a \in S$, so $0 = a + (-a) \in a + S$.

For (2), we prove both implications.

Forward implication. Suppose that $a + S = S$. Since $0 \in S$, we have $a = a + 0 \in a + S = S$.

Reverse implication. Suppose that $a \in S$. If $x \in a + S$, then $x = a + s$ for some $s \in S$, so $x \in S$. Thus $a + S \subseteq S$. Conversely, if $s \in S$, then $s = a + (-a + s) \in a + S$ because $-a + s \in S$. Hence $S \subseteq a + S$.

For (3), translating by $-b$ shows that the equality $a + S = b + S$ is equivalent to $(a - b) + S = S$. By (2), this is equivalent to $a - b \in S$.

For (4), suppose that $(a + S) \cap (b + S) \neq \emptyset$. Then $a + s = b + s'$ for some $s, s' \in S$. Hence $a - b = s' - s \in S$. By (3), $a + S = b + S$. □

Example 8.11 Let $R = \mathbb{Z}$ and $S = 2\mathbb{Z}$. By the division algorithm, for every $a \in R$, there exist $q, r \in \mathbb{Z}$ with $0 \leq r < 2$ such that $a = 2q + r$. Hence

$$a + S = 2q + r + S = r + (2q + S) = r + S.$$

Thus the cosets of $2\mathbb{Z}$ in $\mathbb{Z}$ are $0 + S$ and $1 + S$.

The quotient collapses each row of Figure 5 to a single element: all even integers represent $0 + 2\mathbb{Z}$, while all odd integers represent $1 + 2\mathbb{Z}$.

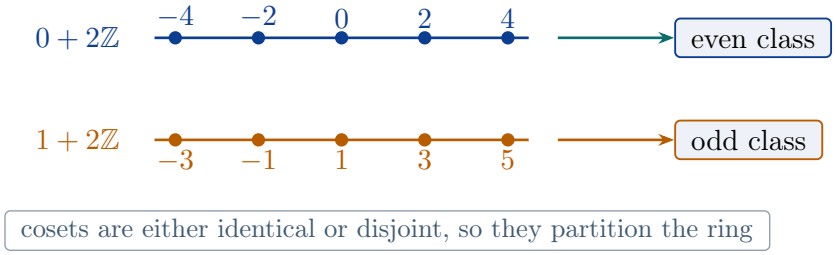


FIGURE 5

Remark 8.12 For $S = 2\mathbb{Z}$, an integer a belongs to $0 + S$ if and only if a is even, and an integer b belongs to $1 + S$ if and only if b is odd.

Example 8.13 Let $R = \mathbb{R}[x]$ and $S = (x)$. By the division algorithm,

$$f(x) = xq(x) + r(x),$$

where $r(x) = 0$ or $\deg(r) < 1$. Hence

$$f(x) + S = xq(x) + r(x) + S = r(x) + S.$$

Thus the coset of $f(x)$ is determined by the constant term of $f(x)$, and the cosets of S in R are

$$\{a + S : a \in \mathbb{R}\}.$$

Lecture 9 — Friday, January 22, 2016

Notation 9.1 Recall that if R is a ring, S is a subring, and $a \in R$, then

$$a + S = \{a + s : s \in S\}.$$

By Proposition 8.10, we have $a + S = b + S$ if and only if $a - b \in S$.

Definition 9.2 Let R be a ring and let S be a subring. We define

$$R/S = \{a + S : a \in R\},$$

read as "R modulo S."

We want operations on R/S to be induced by the operations on R. Define

$$(a + S) + (b + S) = (a + b) + S$$

and

$$(a + S)(b + S) = ab + S.$$

These operations are not always well-defined; the obstruction is exactly whether S is an ideal.

Theorem 9.3 Let R be a ring and let S be a subring of R. Then R/S forms a ring under the operations above if and only if S is an ideal.

Proof. *Forward implication.* Suppose that S is an ideal of R. We first check that the operations are well-defined. Suppose that

$$a + S = \tilde{a} + S \qquad \text{and} \qquad b + S = \tilde{b} + S.$$

By Proposition 8.10, we have $a - \tilde{a} \in S$ and $b - \tilde{b} \in S$. Then

$$(a + b) - (\tilde{a} + \tilde{b}) = (a - \tilde{a}) + (b - \tilde{b}) \in S,$$

so

$$(a + b) + S = (\tilde{a} + \tilde{b}) + S.$$

Thus addition is well-defined.

For multiplication, since S is an ideal, we have $(a - \tilde{a})b \in S$ and $\tilde{a}(b - \tilde{b}) \in S$. Therefore

$$ab - \tilde{a}\tilde{b} = (a - \tilde{a})b + \tilde{a}(b - \tilde{b}) \in S.$$

Hence $ab + S = \tilde{a}\tilde{b} + S$, so multiplication is well-defined. The ring axioms for R/S now follow directly from the ring axioms in R, because the operations are defined by representatives. Therefore R/S is a ring.

Reverse implication. Suppose that R/S is a ring under these operations. Since S is already a subring, it remains to check absorption by elements of R. Let $a \in S$ and $r \in R$. Since $a+S = 0+S$, we have

$$ar + S = (a + S)(r + S) = (0 + S)(r + S) = 0 + S,$$

so $ar \in S$. Similarly,

$$ra + S = (r + S)(a + S) = (r + S)(0 + S) = 0 + S,$$

so $ra \in S$. Hence $S \trianglelefteq R$. □

Remark 9.4 In the construction of quotient rings, ideals play the same role that normal subgroups play in the construction of quotient groups.

Definition 9.5 Any ring of the form R/I, where $I \trianglelefteq R$, is called a **factor ring**.

Remark 9.6

1. If R is commutative and $I \trianglelefteq R$, then
$$(a+I)(b+I) = ab+I = ba+I = (b+I)(a+I),$$
so R/I is commutative.
2. If R has unity 1, then
$$(1+I)(a+I) = 1a+I = a+I.$$
Thus R/I is unital with unity $1+I$.
3. The coset $0+I$ is the additive identity of R/I, since $(0+I)+(a+I) = a+I$ for all $a \in R$.

Example 9.7 Let K be a field. The ring
$$K\langle x, y\rangle$$
is the ring of noncommutative polynomials in x and y over K. For example, $2x^2y \neq 2xyx$ in $K\langle x, y\rangle$. Let
$$I = \{f \in K\langle x, y\rangle : f \text{ has constant term } 0\}.$$
This is a two-sided ideal: sums and additive inverses preserve the zero constant term, and multiplying such a polynomial on either side cannot create a constant term. Modulo I, every polynomial is therefore equal to its constant term, so
$$K\langle x, y\rangle / I \cong K,$$
which is commutative.

Notation 9.8 In R/I, $a+I$ is often written as $\bar{a}$.

Example 9.9 Let $R = \mathbb{Z}$ and $I = n\mathbb{Z}$. Take $\bar{a} \in R/I$. By the division algorithm, $a = nq+r$

for some $q, r \in \mathbb{Z}$ with $0 \leq r < n$. Therefore

$$\bar{a} = \overline{nq + r} = \overline{nq} + \bar{r} = \bar{0} + \bar{r} = \bar{r}.$$

Hence

$$R/I = \{\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{n-1}\}.$$

These elements are distinct, because if $\bar{a} = \bar{b}$ for $a, b \in \{0, 1, \ldots, n-1\}$, then $a - b \in n\mathbb{Z}$, and so $a = b$. Thus $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$.

Remark 9.10

1. $\bar{a} + \bar{b} = \overline{a + b}$.
2. $\bar{a} \cdot \bar{b} = \overline{a \cdot b}$.
3. $\bar{a} = \bar{0}$ if and only if $a \in I$.

Definition 9.11 Let R be a ring and let $a \in R$. The **order** of a, denoted $|a|$, is the smallest positive integer n such that $na = 0$, if such an integer exists.

Example 9.12 In $\mathbb{Z}_{12}$, $|3| = 4$.

Lemma 9.13 Let R be a ring and let $a \in R$. If $m \in \mathbb{N}$ satisfies $ma = 0$, then $|a| \mid m$.

Proof. By the division algorithm, $m = |a|q + r$ with $0 \leq r < |a|$. Then

$$ra = (m - |a|q)a = 0 - 0 = 0.$$

If $r \neq 0$, this contradicts the minimality of $|a|$. Hence $r = 0$, so $|a| \mid m$. □

Example 9.14 We determine the number of elements in $\mathbb{Z}[i]/\langle 2 - i \rangle$. Since $\overline{2 - i} = \bar{0}$, we have $\bar{2} = \bar{i}$. Therefore, for every $a + bi \in \mathbb{Z}[i]$,

$$\overline{a + bi} = \bar{a} + \overline{bi} = \bar{a} + \bar{b}\bar{2} = \overline{a + 2b}.$$

Thus every coset is represented by an integer. Also,

$$\overline{(2-i)(2+i)} = \bar{0},$$

so $\bar{5} = \bar{0}$. Hence

$$\mathbb{Z}[i]/\langle 2-i\rangle = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}\}$$

provided that these five cosets are distinct.

It remains to prove that $|\bar{1}| = 5$. Since $5\bar{1} = \bar{5} = \bar{0}$, Lemma 9.13 gives $|\bar{1}| \mid 5$. Thus $|\bar{1}| \in \{1, 5\}$. Suppose for a contradiction that $|\bar{1}| = 1$. Then $\bar{1} = \bar{0}$, so $1 \in \langle 2-i\rangle$. Hence there exist $a, b \in \mathbb{Z}$ such that

$$1 = (2-i)(a+bi) = (2a+b) + (2b-a)i.$$

Equating real and imaginary parts gives

$$2a + b = 1 \qquad \text{and} \qquad 2b - a = 0.$$

Thus $a = 2b$, and substituting into the first equation gives $5b = 1$, impossible for $b \in \mathbb{Z}$. Therefore $|\bar{1}| \neq 1$, so $|\bar{1}| = 5$. It follows that

$$|\mathbb{Z}[i]/\langle 2-i\rangle| = 5.$$

Lecture 10 — Monday, January 25, 2016

Recall that if R is a ring and $I \trianglelefteq R$, then

$$R/I = \{a + I : a \in R\},$$

with operations

$$(a+I)(b+I) = ab + I \qquad \text{and} \qquad (a+I) + (b+I) = (a+b) + I.$$

These operations are induced by the operations of R. By Theorem 9.3, R/I is a ring if and only if I is an ideal of R. We often write $a + I$ as $\bar{a}$, and Proposition 8.10 gives

$$\bar{a} = \bar{0} \qquad \text{if and only if} \qquad a \in I.$$

The punchline is that the factor ring R/I can be viewed as the ring obtained by setting every element of I equal to zero.

Example 10.1 In $\mathbb{Z}[x]/\langle x^2+1\rangle$, we have

$$\overline{x^3+2x^2+2} = \overline{x^3+2(x^2+1)} = \overline{x^3} = \bar{x}^3.$$

Since $\bar{x}^2 + \bar{1} = \bar{0}$, we have $\bar{x}^2 = -\bar{1}$, and hence

$$\bar{x}^3 = -\bar{x}.$$

Thus $\overline{x^3+2x^2+2} = -\bar{x}$. Also $-\bar{x}^2 = \bar{1}$, so

$$\left(\overline{x^3+2x^2+2}\right)\bar{x} = (-\bar{x})\bar{x} = -\bar{x}^2 = \bar{1}.$$

Therefore the inverse of $\overline{x^3+2x^2+2}$ is $\bar{x}$.

Example 10.2 Let

$$R = \mathbb{Z}_2[x]/\langle x^2+x+1\rangle.$$

We find $(\overline{x+1})^{-1}$. In R,

$$\bar{1} = \overline{1+(x^2+x+1)} = \overline{x^2+x+2} = \overline{x^2+x} = \bar{x}\,\overline{x+1},$$

because $2 = 0$ in $\mathbb{Z}_2$. Hence

$$(\overline{x+1})^{-1} = \bar{x}.$$

Lemma 10.3 Let R be an integral domain and let $f(x), g(x) \in R[x]$ satisfy $\deg(f) = n$ and $\deg(g) = m$. Then $\deg(fg) = m+n$.

Proof. Write

$$f(x) = \sum_{i=0}^{n} a_i x^i \qquad \text{and} \qquad g(x) = \sum_{j=0}^{m} b_j x^j$$

with $a_n, b_m \neq 0$. Then

$$f(x)g(x) = a_n b_m x^{n+m} + \text{terms of lower degree.}$$

Since R is an integral domain and $a_n, b_m \neq 0$, we have $a_n b_m \neq 0$. Thus $\deg(fg) = m+n$. □

Example 10.4 Consider $\mathbb{Z}_4[x]$. We have $\deg(2x+1) = 1$ and $\deg(2x+2) = 1$, but

$$(2x+1)(2x+2) = 2x+2,$$

so

$$\deg((2x+1)(2x+2)) = 1 \neq 1+1.$$

This does not contradict Lemma 10.3, because $\mathbb{Z}_4$ is not an integral domain.

Corollary 10.5 Let R be an integral domain and let $f(x) \in R[x]$ have degree n. If $g(x) \in \langle f(x) \rangle$, then either $g(x) = 0$ or $\deg(g) \geq n$.

Proof. Since $g(x) \in \langle f(x) \rangle$, there exists $h(x) \in R[x]$ such that $g(x) = f(x)h(x)$. If $h(x) = 0$, then $g(x) = 0$. Otherwise, Lemma 10.3 gives

$$\deg(g) = \deg(f) + \deg(h) \geq n.$$

□

Example 10.6 We determine the elements of

$$R = \mathbb{Z}_2[x]/\langle x^2 + x + 1 \rangle.$$

Take $f(x) \in \mathbb{Z}_2[x]$. By the division algorithm,

$$f(x) = g(x)(x^2 + x + 1) + r(x),$$

where $r(x) = 0$ or $\deg(r) < 2$. Hence

$$\overline{f(x)} = \overline{r(x)}.$$

Therefore

$$R = \{\overline{ax+b} : a, b \in \mathbb{Z}_2\}.$$

These elements are distinct. Indeed, suppose that

$$\overline{ax+b} = \overline{cx+d}$$

where $a, b, c, d \in \mathbb{Z}_2$. Then

$$\overline{(a-c)x + (b-d)} = \bar{0},$$

so $(a-c)x + (b-d) \in \langle x^2 + x + 1 \rangle$. By Corollary 10.5, this polynomial must be zero, since it has degree at most 1. Thus $a = c$ and $b = d$. Hence R has $2^2 = 4$ elements.

Theorem 10.7 Let R be commutative with unity. Then $I \trianglelefteq R$ is prime if and only if R/I is an integral domain.

Proof. *Forward implication.* Suppose that I is prime. Since R is commutative with unity, so is R/I; moreover, I is proper, so R/I is nontrivial. Let $\bar{a}, \bar{b} \in R/I$ and suppose that $\bar{a}\bar{b} = \bar{0}$. Then $\overline{ab} = \bar{0}$, so $ab \in I$. Since I is prime, $a \in I$ or $b \in I$. Hence $\bar{a} = \bar{0}$ or $\bar{b} = \bar{0}$. Therefore R/I has no zero divisors, and so R/I is an integral domain.

Reverse implication. Suppose that R/I is an integral domain. Since an integral domain is nontrivial, $I \neq R$. If $ab \in I$, then $\overline{ab} = \bar{0}$, so $\bar{a}\bar{b} = \bar{0}$. Since R/I is an integral domain, $\bar{a} = \bar{0}$ or $\bar{b} = \bar{0}$. Hence $a \in I$ or $b \in I$, and therefore I is prime. □

Theorem 10.8 Let R be commutative and unital. Then $I \trianglelefteq R$ is maximal if and only if R/I is a field.

Proof. *Forward implication.* Suppose that I is maximal. Since I is proper, R/I is nontrivial. Take $0 \neq \bar{a} \in R/I$. Then $a \notin I$. Consider the ideal $\langle a \rangle + I$. We have

$$I \subsetneq \langle a \rangle + I \subseteq R.$$

Since I is maximal, $\langle a \rangle + I = R$. In particular, $1 \in \langle a \rangle + I$, so

$$1 = ab + c$$

for some $b \in R$ and $c \in I$. Passing to cosets gives

$$\bar{1} = \overline{ab + c} = \overline{ab} + \bar{c} = \bar{a}\bar{b} + \bar{0} = \bar{a}\bar{b}.$$

Thus $\bar{a}$ is invertible. Therefore R/I is a field.

Reverse implication. Suppose that R/I is a field. Since a field is nontrivial, $I \neq R$. Let $J \trianglelefteq R$

satisfy

$$I \subsetneq J \subseteq R.$$

We show that $J = R$. Choose $a \in J \setminus I$, so $\bar{a} \neq \bar{0}$ in R/I. Since R/I is a field, there exists $\bar{b} \in R/I$ such that $\bar{a}\bar{b} = \bar{1}$. Thus $ab - 1 \in I \subseteq J$. Since $a \in J$ and J is an ideal, $ab \in J$. Therefore

$$1 = ab - (ab - 1) \in J.$$

Hence $J = R$. Therefore I is maximal. □

Corollary 10.9 Every maximal ideal is prime.

Proof. If I is maximal, then R/I is a field by Theorem 10.8. Every field is an integral domain, so Theorem 10.7 implies that I is prime. □

Example 10.10 Here are some intuitive examples of factor rings that are "isomorphic" to more familiar rings:

1. $\mathbb{Z}[x]/\langle x \rangle \cong \mathbb{Z}$.
2. $\mathbb{Q}[x]/\langle x \rangle \cong \mathbb{Q}$.
3. $\mathbb{Z}/\langle n \rangle = \{\bar{0}, \bar{1}, \ldots, \overline{n-1}\} \cong \mathbb{Z}_n$.

The next goal is to make precise the meaning of isomorphism in these examples.

Lecture 11 — Wednesday, January 27, 2016

Homomorphisms and Isomorphisms

We now make precise the sense in which factor rings such as $\mathbb{Z}/\langle n \rangle$, $\mathbb{Q}[x]/\langle x \rangle$, $\mathbb{Z}[x]/\langle x, 2 \rangle$, and $\mathbb{Z}[x]/\langle x^2 + 1 \rangle$ are the same as the more familiar rings $\mathbb{Z}_n$, $\mathbb{Q}$, $\mathbb{Z}_2$, and $\mathbb{Z}[i]$.

Definition 11.1 Let R and S be rings. A **homomorphism** from R to S is a map $\varphi : R \to S$ such that, for all $a, b \in R$, the following hold:

1. $\varphi(a+b) = \varphi(a) + \varphi(b)$.
2. $\varphi(ab) = \varphi(a)\varphi(b)$.

Example 11.2

1. The map $\varphi : \mathbb{Z} \to \mathbb{Z}_n$ given by $a \mapsto [a]_n$ is a homomorphism. Indeed,
$$\varphi(a+b) = [a+b] = [a] + [b] = \varphi(a) + \varphi(b)$$
and
$$\varphi(ab) = [ab] = [a][b] = \varphi(a)\varphi(b)$$
for all $a, b \in \mathbb{Z}$.
2. The map $\varphi : \mathbb{C} \to \mathbb{R} \oplus \mathbb{R}$ given by $a + bi \mapsto (a, b)$ is not a homomorphism. Since $i^2 = -1$, we have
$$\varphi(i^2) = \varphi(-1) = (-1, 0),$$
but
$$\varphi(i)\varphi(i) = (0,1)(0,1) = (0,0).$$
3. The map $\varphi : M_n(\mathbb{R}) \to \mathbb{R}$ given by $A \mapsto \det(A)$ is not a homomorphism, since $\det(A+B) \neq \det(A) + \det(B)$ in general.

Lemma 11.3 Let $\varphi : R \to S$ be a homomorphism. Then $\varphi(0_R) = 0_S$.

Proof. We have
$$\varphi(0_R) = \varphi(0_R + 0_R) = \varphi(0_R) + \varphi(0_R).$$
Adding $-\varphi(0_R)$ to both sides gives $\varphi(0_R) = 0_S$. □

Definition 11.4 Let $\varphi : R \to S$ be a homomorphism. If φ is bijective, then we call φ an **isomorphism**.

Example 11.5

1. Define $\varphi : \mathbb{R}[x] \to \mathbb{R}$ by $\varphi(f(x)) = f(0)$. This is a homomorphism because

$$\varphi(f(x) + g(x)) = f(0) + g(0) = \varphi(f(x)) + \varphi(g(x))$$

and

$$\varphi(f(x)g(x)) = f(0)g(0) = \varphi(f(x))\varphi(g(x)).$$

It is not an isomorphism, since $\varphi(x + 1) = \varphi(1) = 1$, so φ is not injective.

2. Define $\varphi : \mathbb{Z}_6 \to \mathbb{Z}_2 \oplus \mathbb{Z}_3$ by $\varphi(a) = (a, a)$. First, φ is well-defined: if $[a]_6 = [b]_6$, then $6 \mid (a - b)$, so $2 \mid (a - b)$ and $3 \mid (a - b)$. Hence $[a]_2 = [b]_2$ and $[a]_3 = [b]_3$.

The map φ is a homomorphism, since

$$\varphi(a + b) = (a + b, a + b) = (a, a) + (b, b)$$

and

$$\varphi(ab) = (ab, ab) = (a, a)(b, b).$$

It is also injective. If $\varphi(a) = \varphi(b)$, then $a \equiv b \pmod 2$ and $a \equiv b \pmod 3$. Since $\gcd(2, 3) = 1$, this implies $6 \mid (a - b)$, so $a = b$ in $\mathbb{Z}_6$. Since

$$|\mathbb{Z}_6| = |\mathbb{Z}_2 \oplus \mathbb{Z}_3|,$$

injectivity implies surjectivity. Thus φ is an isomorphism.

Notation 11.6 If R and S are isomorphic, we write $R \cong S$.

If $R \cong S$, then S is a "relabeling" of R with compatible operations.

Example 11.7 Let $R = \{a, b\}$ with operations given by the following tables:

$+$	a	b
a	a	b
b	b	a

$\cdot$	a	b
a	a	a
b	a	b

Then $R \cong \mathbb{Z}_2$.

Lemma 11.8 Let $\varphi : R \to S$ be a homomorphism. Then $\varphi(-r) = -\varphi(r)$ for all $r \in R$.

Proof. We have

$$\varphi(-r) + \varphi(r) = \varphi(-r + r) = \varphi(0_R) = 0_S.$$

Thus $\varphi(-r)$ is the additive inverse of $\varphi(r)$. □

Lemma 11.9 Suppose that R is unital and that $\varphi : R \to S$ is a surjective homomorphism. Then the following hold:

1. The ring S is unital, with $1_S = \varphi(1_R)$.
2. If $a \in \mathcal{U}(R)$, then $\varphi(a) \in \mathcal{U}(S)$ and $\varphi(a)^{-1} = \varphi(a^{-1})$.

Proof. For (1), let $s \in S$. By surjectivity, $s = \varphi(a)$ for some $a \in R$. Then

$$\varphi(1_R)s = \varphi(1_R)\varphi(a) = \varphi(1_R a) = \varphi(a) = s.$$

Similarly, $s\varphi(1_R) = s$. Therefore $\varphi(1_R)$ is the multiplicative identity of S, so $\varphi(1_R) = 1_S$.

For (2), suppose that a^{-1} exists. By (1),

$$\varphi(a^{-1})\varphi(a) = \varphi(a^{-1}a) = \varphi(1_R) = 1_S.$$

Similarly, $\varphi(a)\varphi(a^{-1}) = 1_S$. Therefore $\varphi(a)$ is a unit and $\varphi(a)^{-1} = \varphi(a^{-1})$. □

Theorem 11.10 Let $\varphi : R \to S$ be a homomorphism.

1. For all $r \in R$ and $n \in \mathbb{N}$, $\varphi(r^n) = \varphi(r)^n$ and $\varphi(nr) = n\varphi(r)$.
2. If $A \trianglelefteq R$ and φ is surjective, then $\varphi(A) \trianglelefteq S$.
3. If A is a subring of R, then $\varphi(A)$ is a subring of S.
4. If $B \trianglelefteq S$, then $\varphi^{-1}(B) \trianglelefteq R$, where $\varphi^{-1}(B)$ denotes the preimage of B.

Proof. For (1), both identities follow by induction on n from the defining properties of a homo-

morphism.

For (2), suppose that $A \trianglelefteq R$ and φ is surjective. Since $\varphi(0) = 0$, the set $\varphi(A)$ is nonempty. If $s_1, s_2 \in \varphi(A)$, then $s_1 = \varphi(a)$ and $s_2 = \varphi(b)$ for some $a, b \in A$. Hence

$$s_1 - s_2 = \varphi(a) - \varphi(b) = \varphi(a - b) \in \varphi(A).$$

Now let $s_1 = \varphi(a) \in \varphi(A)$ and let $r \in S$. Since φ is surjective, $r = \varphi(c)$ for some $c \in R$. Then

$$s_1 r = \varphi(a)\varphi(c) = \varphi(ac) \in \varphi(A)$$

and

$$r s_1 = \varphi(c)\varphi(a) = \varphi(ca) \in \varphi(A).$$

By Proposition 5.14, $\varphi(A)$ is an ideal of S.

For (3), if A is a subring of R, then $\varphi(A)$ is nonempty because $\varphi(0) \in \varphi(A)$. If $s_1 = \varphi(a)$ and $s_2 = \varphi(b)$ belong to $\varphi(A)$, then

$$s_1 - s_2 = \varphi(a - b) \in \varphi(A) \qquad \text{and} \qquad s_1 s_2 = \varphi(ab) \in \varphi(A).$$

By Theorem 2.14, $\varphi(A)$ is a subring of S.

For (4), let $B \trianglelefteq S$. The set $\varphi^{-1}(B)$ is nonempty because $\varphi(0_R) = 0_S \in B$. If $a, b \in \varphi^{-1}(B)$, then $\varphi(a), \varphi(b) \in B$, so

$$\varphi(a - b) = \varphi(a) - \varphi(b) \in B.$$

Hence $a - b \in \varphi^{-1}(B)$. If $a \in \varphi^{-1}(B)$ and $r \in R$, then

$$\varphi(ar) = \varphi(a)\varphi(r) \in B \qquad \text{and} \qquad \varphi(ra) = \varphi(r)\varphi(a) \in B.$$

Thus $ar, ra \in \varphi^{-1}(B)$, and Proposition 5.14 shows that $\varphi^{-1}(B) \trianglelefteq R$. □

Example 11.11 The inclusion map $\varphi : \mathbb{Z} \to \mathbb{R}$, defined by $\varphi(a) = a$, is a homomorphism. Although $2\mathbb{Z} \trianglelefteq \mathbb{Z}$, its image is $\varphi(2\mathbb{Z}) = 2\mathbb{Z}$, which is not an ideal of $\mathbb{R}$, since the only ideals of $\mathbb{R}$ are $\{0\}$ and $\mathbb{R}$.

Lecture 12 — Friday, January 29, 2016

Recall that a map $\varphi : R \to S$ is a homomorphism if

$$\varphi(a+b) = \varphi(a) + \varphi(b) \qquad \text{and} \qquad \varphi(ab) = \varphi(a)\varphi(b)$$

for all $a, b \in R$. If φ is bijective, then φ is an isomorphism, and we say that R and S are isomorphic.

Definition 12.1 Let $\varphi : R \to S$ be a homomorphism. The **kernel** of φ, denoted $\ker \varphi$, is

$$\ker \varphi = \{r \in R : \varphi(r) = 0_S\}.$$

The **image** of φ, denoted $\varphi(R)$, is

$$\varphi(R) = \{\varphi(r) : r \in R\} \subseteq S.$$

The map φ is surjective if and only if $\varphi(R) = S$.

Proposition 12.2 Let $\varphi : R \to S$ be a homomorphism. Then $\ker \varphi$ is a two-sided ideal of R; that is, $\ker \varphi \trianglelefteq R$.

Proof. Since $\varphi(0) = 0$, we have $0 \in \ker \varphi$, so $\ker \varphi$ is nonempty. If $a, b \in \ker \varphi$ and $r \in R$, then

$$\varphi(a-b) = \varphi(a) - \varphi(b) = 0 - 0 = 0,$$

so $a - b \in \ker \varphi$. Also,

$$\varphi(ra) = \varphi(r)\varphi(a) = \varphi(r) \cdot 0 = 0$$

and similarly $\varphi(ar) = 0$. By Proposition 5.14, $\ker \varphi \trianglelefteq R$. □

Proposition 12.3 Let $\varphi : R \to S$ be a homomorphism. Then φ is injective if and only if $\ker \varphi = \{0\}$.

Proof. *Forward implication.* Suppose that φ is injective. If $a \in \ker \varphi$, then $\varphi(a) = 0 = \varphi(0)$, so $a = 0$. Hence $\ker \varphi = \{0\}$.

Reverse implication. Suppose that $\ker \varphi = \{0\}$. If $\varphi(a) = \varphi(b)$, then

$$\varphi(a - b) = \varphi(a) - \varphi(b) = 0.$$

Thus $a - b \in \ker \varphi$, so $a - b = 0$, and hence $a = b$. Therefore φ is injective. □

Proposition 12.4

1. The map $\varphi : R \to S$ defined by $\varphi(r) = 0$ for all $r \in R$ is a homomorphism, called the **zero homomorphism**.
2. The map $\varphi : R \to R$ defined by $r \mapsto r$ is the **identity isomorphism** on R. In particular, $R \cong R$.
3. If $\varphi : R \to S$ is an isomorphism, then $\varphi^{-1} : S \to R$ is also an isomorphism. Thus $R \cong S$ implies $S \cong R$.
4. If $\varphi : R \to S$ and $\psi : S \to L$ are isomorphisms, then $\psi \circ \varphi : R \to L$ is an isomorphism. Thus $R \cong S$ and $S \cong L$ imply $R \cong L$.
5. If $\varphi : R \to S$ is an injective homomorphism, then $\varphi : R \to \varphi(R)$ is an isomorphism.

Proof. The claims in (1) and (2) follow directly from the definitions.

For (3), since φ is bijective, φ^{-1} is well-defined. If $s_1, s_2 \in S$, then

$$\varphi^{-1}(s_1 + s_2) = \varphi^{-1}(\varphi(\varphi^{-1}(s_1)) + \varphi(\varphi^{-1}(s_2))) = \varphi^{-1}(\varphi(\varphi^{-1}(s_1) + \varphi^{-1}(s_2))) = \varphi^{-1}(s_1) + \varphi^{-1}(s_2)$$

and

$$\varphi^{-1}(s_1 s_2) = \varphi^{-1}(\varphi(\varphi^{-1}(s_1))\varphi(\varphi^{-1}(s_2))) = \varphi^{-1}(\varphi(\varphi^{-1}(s_1)\varphi^{-1}(s_2))) = \varphi^{-1}(s_1)\varphi^{-1}(s_2).$$

For (4), since φ and ψ are bijective, $\psi \circ \varphi$ is well-defined and bijective. If $a, b \in R$, then

$$\psi \circ \varphi(a + b) = \psi(\varphi(a) + \varphi(b)) = \psi(\varphi(a)) + \psi(\varphi(b)) = \psi \circ \varphi(a) + \psi \circ \varphi(b)$$

and

$$\psi \circ \varphi(ab) = \psi(\varphi(a)\varphi(b)) = \psi(\varphi(a))\psi(\varphi(b)) = \psi \circ \varphi(a)\psi \circ \varphi(b).$$

For (5), since φ is injective, $\varphi : R \to \varphi(R)$ is well-defined and bijective. If $a, b \in R$, then

$$\varphi(a+b) = \varphi(a) + \varphi(b)$$

and

$$\varphi(ab) = \varphi(a)\varphi(b).$$

□

Proposition 12.5 Suppose that $\varphi : R \to S$ is an isomorphism. Then the following hold:

1. If R is commutative, then S is commutative.
2. If R is unital, then S is unital.
3. If $\text{char}(R) = n$, then $\text{char}(S) = n$.
4. If R is an integral domain, then S is an integral domain.
5. If R is a field, then S is a field.

Proof. For (1), take $a', b' \in S$. Since φ is surjective, $a' = \varphi(a)$ and $b' = \varphi(b)$ for some $a, b \in R$. Then

$$a'b' = \varphi(a)\varphi(b) = \varphi(ab) = \varphi(ba) = \varphi(b)\varphi(a) = b'a'.$$

For (2), since R is unital and φ is surjective, Lemma 11.9 shows that S is unital, with unity $\varphi(1_R)$.

For (3), first suppose that $\text{char}(R) = 0$. For a contradiction, assume that $\text{char}(S) = m > 0$. For any $a \in R$,

$$\varphi(ma) = m\varphi(a) = 0.$$

Since φ is injective, $ma = 0$ for all $a \in R$, contradicting $\text{char}(R) = 0$.

Now suppose that $\text{char}(R) = n > 0$. For any $a' \in S$, write $a' = \varphi(a)$. Then

$$na' = n\varphi(a) = \varphi(na) = 0,$$

so $\text{char}(S) = m$ exists and $m \leq n$. Conversely, for any $a \in R$,

$$\varphi(ma) = m\varphi(a) = 0.$$

Since φ is injective, $ma = 0$, so $n \leq m$ by minimality of n. Hence $m = n$.

For (4), by (1) and (2), the ring S is commutative and unital. If $a', b' \in S$ and $a'b' = 0$, write $a' = \varphi(a)$ and $b' = \varphi(b)$. Then

$$0 = a'b' = \varphi(a)\varphi(b) = \varphi(ab).$$

Since φ is injective, $ab = 0$. Because R is an integral domain, $a = 0$ or $b = 0$, and hence $a' = 0$ or $b' = 0$.

For (5), by (1) and (2), the ring S is commutative and unital. If $0 \neq a' \in S$, write $a' = \varphi(a)$. Since φ is injective, $a \neq 0$. Because R is a field, a^{-1} exists, and Lemma 11.9 gives

$$\varphi(a^{-1}) = \varphi(a)^{-1} = (a')^{-1}.$$

Thus every nonzero element of S is a unit. □

Example 12.6 Determine whether the following pairs of rings are isomorphic:

1. $M_2(\mathbb{R})$ and $\mathbb{R}$.
2. $\mathbb{Z}_2[i]$ and $\mathbb{Z}_2$.
3. $\mathbb{Z}$ and $2\mathbb{Z}$.
4. $M_2(\mathbb{Z}_2)$ and $\mathbb{Z}_{16}$.

Solution. For (1), the rings are not isomorphic, since $\mathbb{R}$ is commutative and $M_2(\mathbb{R})$ is not.

For (2), the rings are not isomorphic, since $\mathbb{Z}_2[i]$ has four elements while $\mathbb{Z}_2$ has two.

For (3), the rings are not isomorphic, since $\mathbb{Z}$ is unital and $2\mathbb{Z}$ is not.

For (4), the rings are not isomorphic, since $\text{char}(M_2(\mathbb{Z}_2)) = 2$ but $\text{char}(\mathbb{Z}_{16}) = 16$. □

Proposition 12.7 Let R be a unital ring. Suppose that $\varphi : R \to S$ is an isomorphism and $f(x) \in \mathbb{Z}[x]$. If $r \in R$ satisfies $f(r) = 0$, then $f(\varphi(r)) = 0$.

Lecture 13 — Monday, February 01, 2016

Proof. Write

$$f(x) = \sum_{i=0}^{n} a_i x^i$$

with $a_i \in \mathbb{Z}$. Since φ is an isomorphism of unital rings, $\varphi(1) = 1$. Therefore $\varphi(m) = m$ for every $m \in \mathbb{Z}$, where integers are interpreted as integer multiples of the unity. If $f(r) = 0$, then

$$0 = \varphi(0) = \varphi\left(\sum_{i=0}^{n} a_i r^i\right) = \sum_{i=0}^{n} \varphi(a_i)\varphi(r)^i = \sum_{i=0}^{n} a_i \varphi(r)^i = f(\varphi(r)).$$

□

Example 13.1 We show that $\mathbb{Q}[\sqrt{5}] \ncong \mathbb{Q}[\sqrt{2}]$. Recall that

$$\mathbb{Q}[\sqrt{d}] = \{a + b\sqrt{d} : a, b \in \mathbb{Q}\}.$$

The polynomial $x^2 - 2$ has a root in $\mathbb{Q}[\sqrt{2}]$, namely $\sqrt{2}$. For a contradiction, assume that $\varphi : \mathbb{Q}[\sqrt{2}] \to \mathbb{Q}[\sqrt{5}]$ is an isomorphism. By Proposition 12.7, the polynomial $x^2 - 2$ has a root in $\mathbb{Q}[\sqrt{5}]$. Therefore there exists $a + b\sqrt{5} \in \mathbb{Q}[\sqrt{5}]$ such that

$$(a + b\sqrt{5})^2 - 2 = 0.$$

Expanding gives

$$a^2 + 5b^2 - 2 + 2ab\sqrt{5} = 0.$$

Thus $a^2 + 5b^2 - 2 = 0$ and $2ab = 0$. Since $\mathbb{Q}$ is an integral domain, $a = 0$ or $b = 0$.

1. If $a = 0$, then $5b^2 = 2$, so $b^2 = 2/5$, which has no rational solution.
2. If $b = 0$, then $a^2 = 2$, which has no rational solution.

Both cases give contradictions, so $\mathbb{Q}[\sqrt{2}] \ncong \mathbb{Q}[\sqrt{5}]$.

Definition 13.2 Let R be a ring. An isomorphism $\varphi : R \to R$ is called an **automorphism** of R. The set of automorphisms of R is denoted by $\mathrm{Aut}(R)$.

Proposition 13.3 $\mathrm{Aut}(\mathbb{Z}) = \{\mathrm{Id}_{\mathbb{Z}}\}$.

Proof. Let $\varphi \in \mathrm{Aut}(\mathbb{Z})$. Since φ is a surjective homomorphism between unital rings, Lemma 11.9 gives $\varphi(1) = 1$. It follows that $\varphi(n) = n$ for every positive integer n, and Lemma 11.8 gives $\varphi(-n) = -n$. Therefore $\varphi = \mathrm{Id}_{\mathbb{Z}}$. □

Proposition 13.4 $\mathrm{Aut}(\mathbb{R}) = \{\mathrm{Id}_{\mathbb{R}}\}$.

Proof. Let $\varphi \in \mathrm{Aut}(\mathbb{R})$. Since φ is a surjective homomorphism between unital rings, Lemma 11.9 gives $\varphi(1) = 1$. It follows that $\varphi(n) = n$ for every positive integer n. By Lemma 11.8, homomorphisms preserve additive inverses, so $\varphi(n) = n$ for every integer n. If n is a positive integer, then

$$n\varphi(1/n) = \varphi(n)\varphi(1/n) = \varphi(1) = 1,$$

so $\varphi(1/n) = 1/n$. Therefore $\varphi(q) = q$ for every $q \in \mathbb{Q}$.

We next show that φ preserves the usual order on $\mathbb{R}$. If $a \leq b$, then $b - a \geq 0$, so $b - a = c^2$ for some $c \in \mathbb{R}$. Hence

$$\varphi(b) - \varphi(a) = \varphi(b - a) = \varphi(c^2) = \varphi(c)^2 \geq 0,$$

and therefore $\varphi(a) \leq \varphi(b)$.

Now let $x \in \mathbb{R}$. If $\varphi(x) > x$, choose $q \in \mathbb{Q}$ such that $x < q < \varphi(x)$. Since φ preserves order and fixes $\mathbb{Q}$, we get $\varphi(x) \leq \varphi(q) = q$, contradicting $q < \varphi(x)$. Similarly, if $\varphi(x) < x$, choose $q \in \mathbb{Q}$ such that $\varphi(x) < q < x$. Then $q = \varphi(q) \leq \varphi(x)$, a contradiction. Therefore $\varphi(x) = x$ for every $x \in \mathbb{R}$, so $\varphi = \mathrm{Id}_{\mathbb{R}}$. □

Homomorphisms Between $\mathbb{Z}_n$ and $\mathbb{Z}_m$

Question 13.5 What are all possible homomorphisms $\varphi : \mathbb{Z}_n \to \mathbb{Z}_m$?

Recall that if $a \in R$ and $n \in \mathbb{N}$ satisfy $na = 0$, then $|a| \mid n$.

Lemma 13.6 If $\varphi : R \to S$ is a homomorphism, then $|\varphi(a)| \mid |a|$ for every $a \in R$ with finite order.

Proof. Since

$$|a| \cdot \varphi(a) = \varphi(|a| \cdot a) = \varphi(0) = 0,$$

Lemma 9.13 implies that $|\varphi(a)| \mid |a|$. □

Corollary 13.7 If $\varphi : R \to S$ is an isomorphism, then $|a| = |\varphi(a)|$ for all $a \in R$ with finite order.

Proof. By Lemma 13.6, we have $|\varphi(a)| \mid |a|$. Applying the same lemma to φ^{-1} gives

$$|\varphi^{-1}(\varphi(a))| \mid |\varphi(a)|,$$

so $|a| \mid |\varphi(a)|$. Hence $|a| = |\varphi(a)|$. □

Remark 13.8 If $a \in \mathbb{Z}_n$, then $|a| \mid n$ because $na = 0$.

Example 13.9 The rings $\mathbb{Z}_2 \oplus \mathbb{Z}_4$ and $\mathbb{Z}_8$ are not isomorphic. The ring $\mathbb{Z}_8$ has an element of order 8, namely 1, while every element of $\mathbb{Z}_2 \oplus \mathbb{Z}_4$ has order dividing 4.

Example 13.10 We find all homomorphisms $\varphi : \mathbb{Z}_6 \to \mathbb{Z}_{12}$. Let $\varphi(1) = a \in \mathbb{Z}_{12}$. Then $\varphi(n) = na$ for every $n \in \mathbb{Z}_6$. Since $|\varphi(1)| \mid |1|$, we have $|a| \mid 6$, so $|a| \in \{1, 2, 3, 6\}$. Also, $1 \cdot 1 = 1$ in $\mathbb{Z}_6$, so $a^2 = a$ in $\mathbb{Z}_{12}$. The idempotents in $\mathbb{Z}_{12}$ are

$$0,\ 1,\ 4,\ 9.$$

Their orders are

$$|0| = 1, \qquad |1| = 12, \qquad |4| = 3, \qquad \text{and} \qquad |9| = 4.$$

Thus $a \in \{0, 4\}$.

If $a = 0$, then φ is the zero homomorphism. If $a = 4$, define $\varphi(x) = 4x$. This is well-defined:

if $[u]_6 = [v]_6$, then $6 \mid (u - v)$, so $12 \mid 4(u - v)$, and hence $[4u]_{12} = [4v]_{12}$. For $x, y \in \mathbb{Z}_6$,

$$\varphi(x + y) = 4(x + y) = 4x + 4y$$

and

$$\varphi(xy) = 4xy \equiv 16xy = (4x)(4y) = \varphi(x)\varphi(y) \pmod{12}.$$

Thus there are exactly two homomorphisms.

Many factor rings are isomorphic to more familiar rings. We now formalize this observation.

Example 13.11

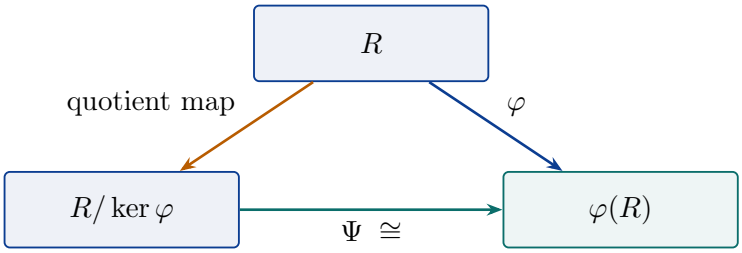
$$\mathbb{Z}[x]/\langle x\rangle \cong \mathbb{Z}, \qquad \mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n, \qquad \text{and} \qquad \mathbb{R}[x]/\langle x^2 + 1\rangle \cong \mathbb{C}.$$

Lecture 14 — Wednesday, February 03, 2016

The First Isomorphism Theorem

Theorem 14.1 (First isomorphism theorem) Let R, S be rings and let $\varphi : R \to S$ be a homomorphism. Then $R/\ker\varphi \cong \varphi(R)$.

The factorization in Figure 6 is the content of the theorem: φ first identifies elements that differ by something in its kernel, and the induced map Ψ then carries the resulting cosets bijectively onto the image.



$\varphi = \Psi \circ \pi$: quotient by the kernel, then identify with the image

FIGURE 6

Proof. Define $\Psi : R/\ker\varphi \to \varphi(R)$ by

$$\Psi(\bar{r}) = \varphi(r).$$

We first check that Ψ is well-defined. Suppose that $\bar{r}_1 = \bar{r}_2$ in $R/\ker\varphi$. Then $r_1 - r_2 \in \ker\varphi$, so $\varphi(r_1 - r_2) = 0$. Hence $\varphi(r_1) = \varphi(r_2)$, and therefore $\Psi(\bar{r}_1) = \Psi(\bar{r}_2)$.

The map Ψ is a homomorphism. If $\bar{a}, \bar{b} \in R/\ker\varphi$, then

$$\Psi(\bar{a} + \bar{b}) = \Psi(\overline{a+b}) = \varphi(a+b) = \varphi(a) + \varphi(b) = \Psi(\bar{a}) + \Psi(\bar{b}),$$

and

$$\Psi(\bar{a}\bar{b}) = \Psi(\overline{ab}) = \varphi(ab) = \varphi(a)\varphi(b) = \Psi(\bar{a})\Psi(\bar{b}).$$

It is surjective by definition of $\varphi(R)$: if $s \in \varphi(R)$, then $s = \varphi(r) = \Psi(\bar{r})$ for some $r \in R$.

Finally, Ψ is injective. If $\bar{a} \in \ker\Psi$, then $\Psi(\bar{a}) = 0$, so $\varphi(a) = 0$. Hence $a \in \ker\varphi$, and therefore $\bar{a} = \bar{0}$ in $R/\ker\varphi$. Thus $\ker\Psi = \{\bar{0}\}$, and Proposition 12.3 shows that Ψ is injective.

Thus Ψ is an isomorphism. □

Corollary 14.2 If $\varphi : R \to S$ is a surjective homomorphism, then $R/\ker\varphi \cong S$.

Proof. If φ is surjective, then $\varphi(R) = S$, so this follows immediately from Theorem 14.1. □

Example 14.3 We show that $\mathbb{Z}/\langle n\rangle \cong \mathbb{Z}_n$. Define $\varphi : \mathbb{Z} \to \mathbb{Z}_n$ by $a \mapsto [a]_n$. This is a homomorphism, and it is surjective because every $[a]_n \in \mathbb{Z}_n$ is equal to $\varphi(a)$.

We claim that $\ker\varphi = \langle n\rangle$. If $a \in \ker\varphi$, then $[a]_n = [0]_n$, so $n \mid a$, and hence $a \in \langle n\rangle$. Conversely, if $a \in \langle n\rangle$, then $a = nk$ for some $k \in \mathbb{Z}$, so

$$\varphi(a) = [nk]_n = [0]_n.$$

Thus $a \in \ker\varphi$, and therefore $\ker\varphi = \langle n\rangle$. By Corollary 14.2,

$$\mathbb{Z}/\langle n\rangle \cong \mathbb{Z}_n.$$

Example 14.4 Let

$$I = \{(a, 0) : a \in R\} \trianglelefteq R \oplus S.$$

We show that $(R \oplus S)/I \cong S$. Define $\varphi : R \oplus S \to S$ by $\varphi(a, b) = b$. This is a homomorphism, and it is surjective because for every $s \in S$, we have $\varphi(0, s) = s$.

The kernel of φ is I. Indeed, if $(a, 0) \in I$, then $\varphi(a, 0) = 0$, so $(a, 0) \in \ker \varphi$. Conversely, if $(a, b) \in \ker \varphi$, then $b = 0$, so $(a, b) = (a, 0) \in I$. Therefore $\ker \varphi = I$, and Corollary 14.2 gives

$$(R \oplus S)/I \cong S.$$

Example 14.5 We show that $\mathbb{Z}[x]/\langle x^2+1\rangle \cong \mathbb{Z}[i]$. Define $\varphi : \mathbb{Z}[x] \to \mathbb{Z}[i]$ by $\varphi(f(x)) = f(i)$. Then

$$\varphi(f(x) + g(x)) = f(i) + g(i) = \varphi(f(x)) + \varphi(g(x)),$$

and similarly $\varphi(f(x)g(x)) = \varphi(f(x))\varphi(g(x))$, so φ is a homomorphism. It is surjective because, for every $a + bi \in \mathbb{Z}[i]$, we have $\varphi(a + bx) = a + bi$.

We claim that $\ker \varphi = \langle x^2 + 1\rangle$. If $f(x) \in \langle x^2 + 1\rangle$, then $f(x) = (x^2 + 1)g(x)$ for some $g(x) \in \mathbb{Z}[x]$, and therefore

$$\varphi(f(x)) = f(i) = (i^2 + 1)g(i) = 0.$$

Thus $\langle x^2 + 1\rangle \subseteq \ker \varphi$.

Conversely, let $f(x) \in \ker \varphi$, so $f(i) = 0$. Because $x^2 + 1$ is monic, polynomial long division works over $\mathbb{Z}$ and gives

$$f(x) = (x^2 + 1)q(x) + r(x),$$

where $r(x) = 0$ or $\deg(r) < 2$. Suppose that $r(x) \neq 0$, and write $r(x) = ax + b$ with $a, b \in \mathbb{Z}$. Then

$$0 = f(i) = (i^2 + 1)q(i) + r(i) = ai + b.$$

Since 1 and i are linearly independent over $\mathbb{Q}$, this forces $a = b = 0$, contradicting $r(x) \neq 0$. Hence $r(x) = 0$, so $f(x) \in \langle x^2 + 1\rangle$. Therefore $\ker \varphi = \langle x^2 + 1\rangle$, and Corollary 14.2 gives

$$\mathbb{Z}[x]/\langle x^2 + 1\rangle \cong \mathbb{Z}[i].$$

Proposition 14.6 Let R be a ring. Then any ideal $I \trianglelefteq R$ is the kernel of some homomorphism $\varphi : R \to S$ for some ring S.

Proof. Define $\varphi : R \to R/I$ by $\varphi(r) = \bar{r}$. This is the **quotient map**, and it is a homomorphism. Moreover,

$$\ker \varphi = \{r \in R : \bar{r} = \bar{0}\} = I.$$

□

Lecture 15 — Friday, February 05, 2016

Fields of Quotients

Theorem 15.1 Let D be an integral domain. Then there exists a field F such that D is isomorphic to a subring of F.

Proof. Let

$$S = \{(a, b) : a, b \in D,\ b \neq 0\}.$$

Define a relation $\sim$ on S by

$$(a, b) \sim (c, d) \qquad \text{if and only if} \qquad ad = bc.$$

This is an equivalence relation. Reflexivity and symmetry are immediate. For transitivity, suppose $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$. Then $ad = bc$ and $cf = de$. Multiplying the first equality by f and the second by b gives

$$adf = bcf = bde.$$

Since $d \neq 0$ and D is an integral domain, cancellation gives $af = be$, so $(a, b) \sim (e, f)$.

Let a/b denote the equivalence class containing $(a, b) \in S$:

$$\frac{a}{b} = \{(c, d) \in S : (a, b) \sim (c, d)\}.$$

Let F be the set of equivalence classes. Define operations on F by

$$\frac{a}{b}+\frac{c}{d}=\frac{ad+bc}{bd} \qquad \text{and} \qquad \frac{a}{b}\cdot\frac{c}{d}=\frac{ac}{bd}.$$

The construction deliberately forgets the chosen representative. In Figure 7, multiplying both coordinates by the same nonzero c changes the pair but not the element of F, because $a(bc)=(ac)b$.

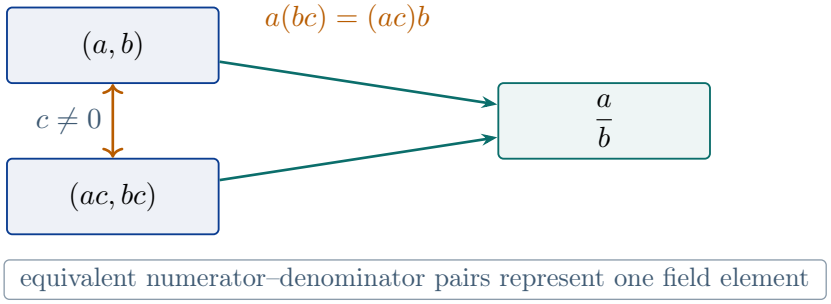


FIGURE 7

Since $b,d\neq 0$ and D is an integral domain, $bd\neq 0$, so these formulas define elements of F.

We check that the operations are well-defined. Suppose $a/b=a'/b'$ and $c/d=c'/d'$. Then $ab'=a'b$ and $cd'=c'd$. For addition,

$$(ad+bc)(b'd')=adb'd'+bcb'd'=(ab')dd'+(cd')bb'=(a'b)dd'+(c'd)bb'=(a'd'+b'c')bd.$$

Thus

$$\frac{ad+bc}{bd}=\frac{a'd'+b'c'}{b'd'}.$$

For multiplication,

$$(ac)(b'd')=(ab')(cd')=(a'b)(c'd)=(a'c')(bd),$$

so

$$\frac{ac}{bd}=\frac{a'c'}{b'd'}.$$

Hence both operations are well-defined.

The ring axioms follow from straightforward calculations using the ring axioms in D. The multiplication is commutative because

$$\frac{a}{b}\cdot\frac{c}{d}=\frac{ac}{bd}=\frac{ca}{db}=\frac{c}{d}\cdot\frac{a}{b}.$$

The unity is $1/1$, since

$$\frac{1}{1} \cdot \frac{a}{b} = \frac{a}{b}.$$

The zero element is $0/1$, and $a/b = 0/1$ if and only if $a = 0$. If $a/b \neq 0$, then $a \neq 0$, so $(b, a) \in S$ and

$$\frac{a}{b} \cdot \frac{b}{a} = \frac{ab}{ab} = \frac{1}{1}.$$

Thus every nonzero element of F is invertible, and F is a field.

Now define $\rho : D \to F$ by $\rho(a) = a/1$. Then

$$\rho(a + b) = \frac{a + b}{1} = \frac{a}{1} + \frac{b}{1} = \rho(a) + \rho(b)$$

and

$$\rho(ab) = \frac{ab}{1} = \frac{a}{1} \cdot \frac{b}{1} = \rho(a)\rho(b),$$

so ρ is a homomorphism. Moreover,

$$\rho(a) = 0/1 \qquad \text{if and only if} \qquad (a, 1) \sim (0, 1) \qquad \text{if and only if} \qquad a = 0.$$

Thus $\ker \rho = \{0\}$, so ρ is injective. Therefore $D \cong \rho(D)$, and $\rho(D)$ is a subring of F. □

Definition 15.2 The field F constructed above is called the **field of quotients**, or **field of fractions**, of D.

Example 15.3 If $D = \mathbb{Z}$, then its field of quotients is isomorphic to $\mathbb{Q}$. If $D = K[x]$, then its field of quotients is

$$\left\{ \frac{f(x)}{g(x)} : f(x), g(x) \in K[x],\ g(x) \neq 0 \right\},$$

the field of rational functions over K.

Notation 15.4 If K is a field, we denote the field of rational functions over K by

$$K(x) = \left\{ \frac{f(x)}{g(x)} : f(x), g(x) \in K[x],\ g(x) \neq 0 \right\}.$$

Example 15.5 An infinite field of characteristic $p > 0$ is $\mathbb{Z}_p(x)$.

Lecture 16 — Monday, February 08, 2016

3. Polynomial Rings and Factorization

Polynomial Rings and Roots

If D is an integral domain and $f(x), g(x) \in D[x]$ are nonzero, then

$$\deg(f(x)g(x)) = \deg(f(x)) + \deg(g(x)).$$

Let F be a field, and recall the division algorithm for $F[x]$ (Theorem 7.9): if $f(x), g(x) \in F[x]$ with $g(x) \neq 0$, then there exist unique polynomials $q(x), r(x) \in F[x]$ such that

$$f(x) = g(x)q(x) + r(x),$$

where either $r(x) = 0$ or $\deg(r) < \deg(g)$.

Example 16.1 We find the remainder when $3x^4 + x^3 + 2x^2$ is divided by $x^2 + 4x + 2$ in $\mathbb{Z}_5[x]$.

$$\begin{array}{r} 3x^2 + 4x + 0 \\ x^2 + 4x + 2 \,\big)\, 3x^4 + x^3 + 2x^2 + 0x + 0 \\ \underline{3x^4 + 2x^3 + x^2} \quad \downarrow \\ 4x^3 + x^2 + 0x \\ \underline{4x^3 + x^2 + 3x} \quad \downarrow \\ 2x + 0 \end{array}$$

Thus

$$3x^4 + x^3 + 2x^2 = (3x^2 + 4x)(x^2 + 4x + 2) + 2x$$

in $\mathbb{Z}_5[x]$, so the remainder is $2x$.

Definition 16.2 Let R be a ring, and let $f(x), g(x) \in R[x]$. We say that $g(x)$ **divides** $f(x)$, and write $g(x) \mid f(x)$, if there exists a polynomial $h(x) \in R[x]$ such that

$$f(x) = g(x)h(x).$$

Definition 16.3 Let R be a ring and let $f(x) \in R[x]$. An element $a \in R$ is a **root**, or **zero**, of $f(x)$ if $f(a) = 0$.

Remark 16.4 If R is a ring and $f(x) \in R[x]$, then $f(x)$ determines a function $f : R \to R$ by evaluation. However, two distinct polynomials may determine the same function.

Example 16.5 In $\mathbb{Z}_p[x]$, the polynomials $f(x) = 0$ and $g(x) = x^p - x$ are distinct. Nevertheless, by Fermat's little theorem, $g(a) = 0$ for every $a \in \mathbb{Z}_p$. Thus they define the same function $\mathbb{Z}_p \to \mathbb{Z}_p$.

Proposition 16.6 Let F be a field, let $a \in F$, and let $f(x) \in F[x]$. Then $f(a)$ is the remainder when $f(x)$ is divided by $x - a$.

Proof. By Theorem 7.9, there exist $q(x), r(x) \in F[x]$ such that

$$f(x) = (x - a)q(x) + r(x),$$

where either $r(x) = 0$ or $\deg(r) < 1$. Hence $r(x)$ is constant, and evaluating at $x = a$ gives

$$f(a) = (a - a)q(a) + r = r.$$

Therefore the remainder is $f(a)$. □

Proposition 16.7 Let F be a field and let $f(x) \in F[x]$. Then $a \in F$ is a root of $f(x)$ if and only if $(x - a) \mid f(x)$.

Proof. *Forward implication.* Suppose that a is a root of $f(x)$. By Proposition 16.6, the remainder when $f(x)$ is divided by $x - a$ is $f(a) = 0$. Hence $f(x) = (x - a)q(x)$ for some $q(x) \in F[x]$, so $(x - a) \mid f(x)$.

Reverse implication. Suppose that $(x-a) \mid f(x)$. Then $f(x) = (x-a)g(x)$ for some $g(x) \in F[x]$. Evaluating at a gives

$$f(a) = (a-a)g(a) = 0,$$

so a is a root of $f(x)$. □

Combining the remainder theorem and the factor theorem gives the two equivalent routes in Figure 8. Evaluation at a extracts the remainder modulo $x-a$, and that remainder vanishes exactly when $x-a$ is a factor.

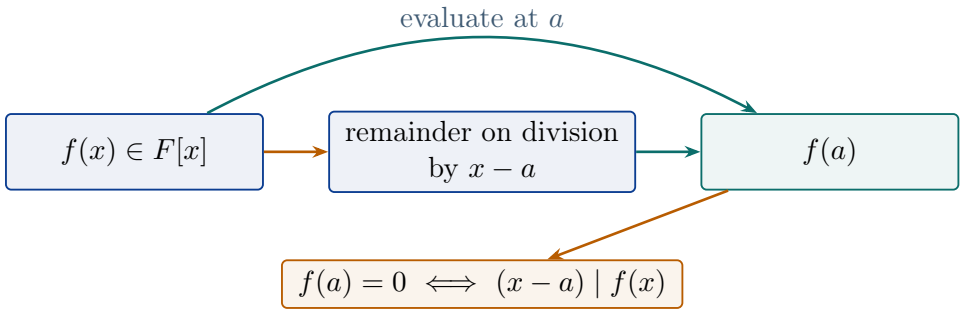


FIGURE 8

By induction, if $a_1, \ldots, a_n$ are distinct roots of $f(x) \in F[x]$, then

$$(x-a_1)(x-a_2)\cdots(x-a_n) \mid f(x).$$

Definition 16.8 Suppose that F is a field and $f(x) \in F[x]$. If $k \in \mathbb{N}$, $(x-a)^k \mid f(x)$, and $(x-a)^{k+1} \nmid f(x)$, then $a \in F$ is a **root of $f(x)$ of multiplicity k**.

Proposition 16.9 Let F be a field and let $f(x) \in F[x]$ with $\deg(f) = n$. Then $f(x)$ has at most n roots in F, counted with multiplicity.

Proof. Suppose that $f(x)$ has m roots counted with multiplicity, where $m > n$. Write these roots as $a_1, \ldots, a_m \in F$, allowing repetitions according to multiplicity. Then

$$f(x) = (x-a_1)\cdots(x-a_m)g(x)$$

for some nonzero $g(x) \in F[x]$. Therefore

$$n = \deg(f) = m + \deg(g) \geq m > n,$$

a contradiction. □

It is natural to ask which properties pass from a ring R to the polynomial ring $R[x]$. Commutativity, existence of unity, characteristic, and the absence of zero divisors behave well. However, being a field does not pass to polynomial rings: $\mathbb{R}$ is a field, but $\mathbb{R}[x]$ is not. Similarly, the property that every ideal is principal does not pass from $\mathbb{Z}$ to $\mathbb{Z}[x]$; for instance, $\langle x, 2\rangle$ is not principal in $\mathbb{Z}[x]$.

Proposition 16.10 If D is an integral domain, then $D[x]$ is an integral domain.

Proof. The ring $D[x]$ is commutative and unital because D is. It remains to show that $D[x]$ has no zero divisors. Let

$$f(x) = \sum_{i=0}^{n} a_i x^i \qquad \text{and} \qquad g(x) = \sum_{j=0}^{m} b_j x^j$$

be nonzero polynomials, with $a_n \neq 0$ and $b_m \neq 0$. The leading term of $f(x)g(x)$ is $a_n b_m x^{n+m}$. Since D is an integral domain, $a_n b_m \neq 0$. Hence $f(x)g(x) \neq 0$, so $D[x]$ has no zero divisors. □

Lecture 17 — Wednesday, February 10, 2016

We now apply the root criterion from Proposition 16.7 and the root bound from Proposition 16.9.

Example 17.1 Let K be a field. For $a \in K$, we have

$$K[x]/\langle x - a\rangle \cong K.$$

Proof. Define $\varphi : K[x] \to K$ by $\varphi(f(x)) = f(a)$. This is a homomorphism, and it is surjective because every $c \in K$ is the image of the constant polynomial c. Moreover,

$$f(x) \in \ker \varphi$$

if and only if $f(a) = 0$, which holds if and only if $(x - a) \mid f(x)$ by Proposition 16.7. Thus

$\ker \varphi = \langle x - a \rangle$. By the first isomorphism theorem, Theorem 14.1,

$$K[x]/\langle x - a \rangle \cong K.$$

□

Remark 17.2 The analogous statement over $\mathbb{C}[x]$ says that the kernel of the evaluation map $\mathbb{C}[x] \to \mathbb{C}$ at i is $\langle x - i \rangle$. For the evaluation map $\mathbb{Z}[x] \to \mathbb{Z}[i]$, the corresponding kernel is $\langle x^2 + 1 \rangle$, as in Example 14.5.

Example 17.3 There can be more roots than the degree when the coefficient ring is not an integral domain. In $\mathbb{Z}_8[x]$, the polynomial $x^2 - 1$ has roots $1, 3, 5$, and 7. Indeed,

$$1^2 \equiv 3^2 \equiv 5^2 \equiv 7^2 \equiv 1 \pmod 8.$$

Thus $x^2 - 1$ has four roots in $\mathbb{Z}_8$, even though it has degree 2.

Proposition 17.4 If D is an integral domain, then $\mathcal{U}(D[x]) = \mathcal{U}(D)$.

Solution. First, every unit of D is a unit of $D[x]$, since a constant polynomial whose inverse lies in D has the same inverse when regarded as an element of $D[x]$.

Conversely, let $f(x) \in \mathcal{U}(D[x])$. Then there exists $g(x) \in D[x]$ such that

$$f(x)g(x) = 1.$$

Since D is an integral domain, degrees add for nonzero polynomials in $D[x]$. Therefore

$$\deg(f) + \deg(g) = \deg(1) = 0.$$

Thus $\deg(f) = \deg(g) = 0$, so f and g are constant polynomials. Hence $f \in D$ and f has inverse $g \in D$, so $f \in \mathcal{U}(D)$. Therefore $\mathcal{U}(D[x]) = \mathcal{U}(D)$. □

Example 17.5 Show that $2x + 1$ is invertible in $\mathbb{Z}_4[x]$.

Solution. We have

$$(2x+1)(2x+1) = 4x^2 + 4x + 1 = 1$$

in $\mathbb{Z}_4[x]$. Hence $2x+1$ is its own inverse. □

Example 17.6 Let F be an infinite field. Show that if $f(a) = 0$ for infinitely many $a \in F$, then $f(x) = 0$ in $F[x]$.

Solution. If $f(x) \neq 0$, then Proposition 16.9 shows that $f(x)$ has at most $\deg(f)$ roots in F. This contradicts the assumption that $f(x)$ has infinitely many roots. Hence $f(x) = 0$. □

Corollary 17.7 If F is an infinite field and $f(x), g(x) \in F[x]$ agree at infinitely many values of F, then $f(x) = g(x)$.

Proof. Apply Example 17.6 to $h(x) = f(x) - g(x)$. □

Example 17.8 We find a polynomial in $\mathbb{Z}[x]$ whose roots include $\frac{1}{2}$ and $-\frac{1}{3}$.

Solution. In $\mathbb{Q}[x]$, the polynomial

$$\left(x - \frac{1}{2}\right)\left(x + \frac{1}{3}\right)$$

has the desired roots. Multiplying by 6 gives the integer polynomial

$$6\left(x - \frac{1}{2}\right)\left(x + \frac{1}{3}\right) = 6x^2 - x - 1.$$

□

Theorem 17.9 (Fundamental theorem of algebra) Let $f(x) \in \mathbb{C}[x]$ with $\deg(f) \geq 1$. Then $f(x)$ has a root in $\mathbb{C}$.

We borrow this theorem from PMATH 352 without proof.

Corollary 17.10 Let $f(x) \in \mathbb{C}[x]$ with $\deg(f) = n \geq 1$. Then

$$f(x) = \alpha(x - a_1)(x - a_2) \cdots (x - a_n)$$

for some $\alpha, a_1, \ldots, a_n \in \mathbb{C}$.

Proof. We argue by induction on $\deg(f)$. If $\deg(f) = 1$, the result is immediate. If $\deg(f) \geq 2$, then Theorem 17.9 gives a root $a_1 \in \mathbb{C}$, so $f(x) = (x - a_1)g(x)$ for some $g(x) \in \mathbb{C}[x]$. Applying the induction hypothesis to $g(x)$ gives the desired factorization. □

Proposition 17.11 Let $f(x) \in \mathbb{R}[x]$. If $a \in \mathbb{C}$ is a root of $f(x)$, then $\overline{a}$ is also a root of $f(x)$.

Proof. Write $f(x) = \sum_{i=0}^{n} b_i x^i$, where $b_i \in \mathbb{R}$. Since $f(a) = 0$,

$$\sum_{i=0}^{n} b_i a^i = 0.$$

Therefore

$$f(\overline{a}) = \sum_{i=0}^{n} b_i \overline{a}^{\,i} = \sum_{i=0}^{n} \overline{b_i a^i} = \overline{\sum_{i=0}^{n} b_i a^i} = 0.$$

□

Corollary 17.12 If $f(x) \in \mathbb{R}[x]$ and $\deg(f) \geq 1$, then $f(x)$ can be factored over $\mathbb{R}$ into linear factors and irreducible monic quadratic factors.

Proof. By Corollary 17.10, $f(x)$ factors into linear factors over $\mathbb{C}$. The nonreal roots occur in conjugate pairs by Proposition 17.11. If $\beta \in \mathbb{C} \setminus \mathbb{R}$ is a root, then

$$(x - \beta)(x - \overline{\beta}) = x^2 - (\beta + \overline{\beta})x + \beta\overline{\beta} = x^2 - 2\,\mathrm{Re}(\beta)x + |\beta|^2 \in \mathbb{R}[x].$$

Thus, apart from the nonzero leading constant, the real roots give linear factors over $\mathbb{R}$, and the nonreal conjugate pairs give irreducible monic quadratic factors over $\mathbb{R}$. □

Irreducibility of Polynomials

Definition 17.13 Let D be an integral domain. A polynomial $f(x) \in D[x]$ is **irreducible** over D if the following conditions hold:

1. $f(x) \neq 0$ and $f(x) \notin \mathcal{U}(D[x])$.
2. Whenever $f(x) = g(x)h(x)$ with $g(x), h(x) \in D[x]$, at least one of $g(x)$ and $h(x)$ is a unit in $D[x]$.

If $f(x)$ is a nonzero nonunit that is not irreducible, then it is **reducible**.

Lecture 18 — Friday, February 12, 2016

Example 18.1 The polynomial $2x + 2$ is irreducible over $\mathbb{Q}$. Indeed, if

$$2x + 2 = g(x)h(x)$$

with $g(x), h(x) \in \mathbb{Q}[x]$, then $g(x)$ and $h(x)$ are nonzero and

$$1 = \deg(2x + 2) = \deg(g) + \deg(h).$$

Hence either $\deg(g) = 0$ or $\deg(h) = 0$, so one factor is a unit in $\mathbb{Q}[x]$.

Example 18.2 The same polynomial is reducible over $\mathbb{Z}$, since

$$2x + 2 = 2(x + 1)$$

and neither 2 nor $x + 1$ is a unit in $\mathbb{Z}[x]$.

Example 18.3 The polynomial $x^2 + 1$ is irreducible over $\mathbb{R}$, but reducible over $\mathbb{C}$.

Example 18.4 The polynomial $x^2 - 2$ is irreducible over $\mathbb{Q}$, but reducible over $\mathbb{R}$.

Question 18.5 Which polynomials in $\mathbb{C}[x]$ are irreducible over $\mathbb{C}$?

By the fundamental theorem of algebra, every polynomial in $\mathbb{C}[x]$ of degree at least 2 factors into linear terms. The nonzero constant polynomials are units. Therefore the irreducible polynomials over $\mathbb{C}$ are precisely

$$\{ax + b : a, b \in \mathbb{C},\ a \neq 0\}.$$

Let $K \subseteq F$ be fields. Observe that if $f(x) \in K[x]$ is irreducible over F, then $f(x)$ is irreducible over K. Our goal is to develop tests for deciding when a polynomial $f(x) \in D[x]$, with D an integral domain, is irreducible.

Lemma 18.6 Let F be a field. If $f(x) \in F[x]$ has degree at least 2 and has a root in F, then $f(x)$ is reducible.

Proof. Let $a \in F$ be a root. By Proposition 16.7, $f(x) = (x-a)g(x)$ for some $g(x) \in F[x]$. Since $\deg(f) \geq 2$, we have $\deg(g) \geq 1$, so neither factor is a unit. Thus $f(x)$ is reducible. □

Remark 18.7 The converse is false. The polynomial

$$x^4 + 2x^2 + 1 = (x^2 + 1)^2$$

has no roots in $\mathbb{R}$, but it is reducible over $\mathbb{R}$.

Proposition 18.8 (Reducibility test for degree 2 or 3) Let F be a field. If $f(x) \in F[x]$ has degree 2 or 3, then $f(x)$ is reducible over F if and only if $f(x)$ has a root in F.

Proof. *Forward implication.* Suppose that $f(x)$ is reducible. Then $f(x) = g(x)h(x)$ for some nonunits $g(x), h(x) \in F[x]$. Hence $\deg(g) \geq 1$ and $\deg(h) \geq 1$. Since $\deg(f) \in \{2, 3\}$, one of g and h has degree 1. Without loss of generality, say $\deg(g) = 1$, so $g(x) = ax + b$ with $a \neq 0$. Then $g(-a^{-1}b) = 0$, so $f(x)$ has a root in F.

Reverse implication. This is Lemma 18.6. □

Example 18.9 Determine whether the following polynomials are irreducible.

1. In $\mathbb{Z}_3[x]$, determine whether x^2+1 is irreducible.
2. In $\mathbb{Z}_5[x]$, determine whether x^2+1 is irreducible.
3. In $\mathbb{Z}_3[x]$, determine whether x^3+x^2+x+1 is irreducible.
4. In $\mathbb{Z}_p[x]$, where p is prime, determine whether
$$x^{p-1}+x^{p-2}+\cdots+x+1$$
is irreducible.

Solution. For (1), the polynomial x^2+1 has no roots in $\mathbb{Z}_3$. Since it has degree 2, Proposition 18.8 shows that it is irreducible over $\mathbb{Z}_3$.

For (2), the polynomial x^2+1 is reducible because 2 is a root.

For (3), the polynomial x^3+x^2+x+1 is reducible because 2 is a root.

For (4), if $p=2$, then the polynomial is $x+1$, which is irreducible over $\mathbb{Z}_2$. If $p>2$, then 1 is a root, since
$$1^{p-1}+1^{p-2}+\cdots+1+1=p=0$$
in $\mathbb{Z}_p$. Therefore the polynomial is reducible for odd primes p. □

Example 18.10 The polynomial x^2-2 is irreducible over $\mathbb{Q}$ because it has no roots in $\mathbb{Q}$ and has degree 2.

Example 18.11 The polynomial x^2-2 is irreducible over $\mathbb{Z}$ but reducible over $\mathbb{R}$, even though $\mathbb{Z}\subseteq\mathbb{R}$. Thus irreducibility can be lost when the coefficient ring is enlarged.

Question 18.12 If $f(x)\in\mathbb{Z}[x]$ is reducible over $\mathbb{Q}$, then is it reducible over $\mathbb{Z}$?

Yes. But we cannot yet prove it.

Definition 18.13 Let

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{Z}[x].$$

The **content** of $f(x)$ is

$$\text{content}(f) = \gcd(a_0, a_1, \ldots, a_n).$$

Definition 18.14 A polynomial $f(x) \in \mathbb{Z}[x]$ is **primitive** if $\text{content}(f) = 1$.

Lemma 18.15 The product of two primitive polynomials in $\mathbb{Z}[x]$ is primitive.

Proof. Suppose that $f(x), g(x) \in \mathbb{Z}[x]$ are primitive. For a contradiction, suppose that $f(x)g(x)$ is not primitive. Then some prime p divides every coefficient of $f(x)g(x)$.

For $h(x) \in \mathbb{Z}[x]$, let $\overline{h}(x) \in \mathbb{Z}_p[x]$ denote the polynomial obtained by reducing the coefficients of $h(x)$ modulo p. Since p divides every coefficient of $f(x)g(x)$, we have

$$\overline{f}(x)\overline{g}(x) = \overline{0}$$

in $\mathbb{Z}_p[x]$. By Proposition 16.10, $\mathbb{Z}_p[x]$ is an integral domain, so $\overline{f}(x) = \overline{0}$ or $\overline{g}(x) = \overline{0}$. Thus p divides every coefficient of $f(x)$ or every coefficient of $g(x)$, contradicting that both polynomials are primitive. Therefore $f(x)g(x)$ is primitive. □

Lecture 19 — Wednesday, February 24, 2016

Gauss's Lemma and Irreducibility Tests

We will use Lemma 18.15: the product of two primitive polynomials in $\mathbb{Z}[x]$ is primitive.

Theorem 19.1 (Gauss's lemma) Let $f(x) \in \mathbb{Z}[x]$. If $f(x)$ is reducible over $\mathbb{Q}$, then $f(x)$ is reducible over $\mathbb{Z}$.

Proof. Suppose that $f(x)$ is reducible over $\mathbb{Q}$. Then

$$f(x) = g(x)h(x)$$

for some nonconstant $g(x), h(x) \in \mathbb{Q}[x]$.

First suppose that $f(x)$ is not primitive. Let $c = \text{content}(f)$. Then $c \neq 1$ and $f(x) = cf_1(x)$ for some $f_1(x) \in \mathbb{Z}[x]$, so $f(x)$ is reducible over $\mathbb{Z}$.

Now suppose that $f(x)$ is primitive. Choose positive integers a and b clearing the denominators of the coefficients of $g(x)$ and $h(x)$, respectively, so that $ag(x), bh(x) \in \mathbb{Z}[x]$. Let

$$c_1 = \text{content}(ag) \qquad \text{and} \qquad c_2 = \text{content}(bh).$$

Then

$$ag(x) = c_1 g_1(x) \qquad \text{and} \qquad bh(x) = c_2 h_1(x),$$

where $g_1(x), h_1(x) \in \mathbb{Z}[x]$ are primitive. Hence

$$abf(x) = c_1 c_2 g_1(x) h_1(x).$$

Taking contents on both sides gives $ab = c_1 c_2$, because $f(x)$ is primitive and $g_1(x)h_1(x)$ is primitive by Lemma 18.15. Therefore

$$f(x) = g_1(x)h_1(x),$$

and this is a factorization in $\mathbb{Z}[x]$ into nonconstant polynomials. Thus $f(x)$ is reducible over $\mathbb{Z}$. □

Remark 19.2 The proof shows a little more. If $f(x) \in \mathbb{Z}[x]$ and $f(x) = g(x)h(x)$ with $g(x), h(x) \in \mathbb{Q}[x]$, then there exist $g_1(x), h_1(x) \in \mathbb{Z}[x]$ such that

$$f(x) = g_1(x)h_1(x),$$

with $\deg(g_1) = \deg(g)$ and $\deg(h_1) = \deg(h)$.

Remark 19.3 The contrapositive of Theorem 19.1 says that if $f(x) \in \mathbb{Z}[x]$ is irreducible over $\mathbb{Z}$, then $f(x)$ is irreducible over $\mathbb{Q}$. That is definitely true. However, the following converses are false:

1. If $f(x)$ is reducible over $\mathbb{Z}$, then $f(x)$ is reducible over $\mathbb{Q}$.

2. If $f(x)$ is irreducible over $\mathbb{Q}$, then $f(x)$ is irreducible over $\mathbb{Z}$.

For example, $2x + 2$ is reducible over $\mathbb{Z}$ but irreducible over $\mathbb{Q}$.

The following theorem plays nicely with Gauss's lemma to give a test for irreducibility over $\mathbb{Q}$.

Theorem 19.4 (Rational roots theorem) Let

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{Z}[x],$$

where $a_n \neq 0$ and $a_0 \neq 0$. If $(p/q) \in \mathbb{Q}$, $\gcd(p, q) = 1$, and $f(p/q) = 0$, then $p \mid a_0$ and $q \mid a_n$.

Proof. From $f(p/q) = 0$, multiplying by q^n gives

$$a_n p^n + a_{n-1} p^{n-1} q + \cdots + a_1 p q^{n-1} + a_0 q^n = 0.$$

Rearranging shows that $p \mid a_0 q^n$. Since $\gcd(p, q) = 1$, this implies $p \mid a_0$. Similarly, rearranging the same equation shows that $q \mid a_n p^n$, and hence $q \mid a_n$. □

The rational roots theorem gives *candidates* for rational roots. The conditions $p \mid a_0$ and $q \mid a_n$ are necessary, not sufficient. For instance, the candidates for rational roots of $x^2 + 1$ are ± 1, but neither is a root.

Example 19.5 Determine if the following polynomials are irreducible over $\mathbb{Q}$:

1. $f(x) = x^2 + 7x + 1$.
2. $f(x) = x^3 + x^2 + 2$.

Solution. For (1), the quadratic formula gives roots

$$\frac{-7 \pm \sqrt{45}}{2},$$

which are not rational. Since $\deg(f) = 2$, Proposition 18.8 shows that $f(x)$ is irreducible over $\mathbb{Q}$.

For (2), the rational root candidates are ± 1 and ± 2. None of these is a root, so $f(x)$ has no root

in $\mathbb{Q}$. Since $\deg(f) = 3$, Proposition 18.8 shows that $f(x)$ is irreducible over $\mathbb{Q}$. □

We now develop irreducibility tests that can be useful beyond degrees 2 and 3.

Theorem 19.6 (Mod p irreducibility test) Let p be prime and let $f(x) \in \mathbb{Z}[x]$ have degree greater than 1. Let $\overline{f}(x) \in \mathbb{Z}_p[x]$ be obtained by reducing the coefficients of $f(x)$ modulo p. If $\overline{f}(x)$ is irreducible over $\mathbb{Z}_p$ and $\deg(\overline{f}) = \deg(f)$, then $f(x)$ is irreducible over $\mathbb{Q}$.

Proof. Suppose that $f(x)$ is reducible over $\mathbb{Q}$. By Gauss's lemma, Theorem 19.1, and Remark 19.2, there exist nonconstant polynomials $g(x), h(x) \in \mathbb{Z}[x]$ such that

$$f(x) = g(x)h(x).$$

Reducing modulo p gives

$$\overline{f}(x) = \overline{g}(x)\overline{h}(x)$$

in $\mathbb{Z}_p[x]$. Since $\overline{f}(x)$ is irreducible over $\mathbb{Z}_p$, one of $\overline{g}(x)$ and $\overline{h}(x)$ is a unit. Suppose, without loss of generality, that $\overline{g}(x)$ is a unit. Then $\deg(\overline{g}) = 0$, so

$$\deg(\overline{f}) = \deg(\overline{h}) \leq \deg(h) < \deg(f),$$

contradicting $\deg(\overline{f}) = \deg(f)$. Therefore $f(x)$ is irreducible over $\mathbb{Q}$. □

Lecture 20 — Friday, February 26, 2016

We continue using the mod p irreducibility test, Theorem 19.6.

Example 20.1 Determine if the following polynomials are irreducible over $\mathbb{Q}$.

1. $$f(x) = 7x^3 - 6x^2 + 3x + 9.$$

2.
$$f(x) = \frac{3}{7}x^4 - \frac{2}{7}x^2 + \frac{8}{35}x + \frac{3}{5}.$$

Solution. For (1), reducing modulo 2 gives

$$\overline{f}(x) = x^3 + x + 1 \in \mathbb{Z}_2[x].$$

Since $\overline{f}(0) = 1$ and $\overline{f}(1) = 1$, this cubic has no roots in $\mathbb{Z}_2$. By Proposition 18.8, $\overline{f}(x)$ is irreducible over $\mathbb{Z}_2$. Also $\deg(\overline{f}) = \deg(f)$, so Theorem 19.6 shows that $f(x)$ is irreducible over $\mathbb{Q}$.

For (2), multiplying by the nonzero rational number 35 does not affect irreducibility over $\mathbb{Q}$, so set

$$h(x) = 35f(x) = 15x^4 - 10x^2 + 8x + 21 \in \mathbb{Z}[x].$$

Reducing modulo 2 gives

$$\overline{h}(x) = x^4 + x + 1 \in \mathbb{Z}_2[x].$$

This polynomial has no roots in $\mathbb{Z}_2$. Since it has degree 4, the only possible nontrivial factorization over $\mathbb{Z}_2$ would be a product of two irreducible quadratics. The only irreducible quadratic in $\mathbb{Z}_2[x]$ is $x^2 + x + 1$, and

$$(x^2 + x + 1)^2 = x^4 + x^2 + 1 \neq x^4 + x + 1.$$

Thus $\overline{h}(x)$ is irreducible over $\mathbb{Z}_2$. By Theorem 19.6, $h(x)$ is irreducible over $\mathbb{Q}$, and hence so is $f(x)$. □

Theorem 20.2 (Eisenstein's criterion) Let

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{Z}[x].$$

If there exists a prime p such that $p \nmid a_n$, $p \mid a_i$ for $0 \leq i \leq n-1$, and $p^2 \nmid a_0$, then $f(x)$ is irreducible over $\mathbb{Q}$.

Proof. For a contradiction, suppose that $f(x)$ is reducible over $\mathbb{Q}$. By the proof of Gauss's lemma, we may write

$$f(x) = g(x)h(x),$$

where $g(x), h(x) \in \mathbb{Z}[x]$ and both have positive degree. Write

$$g(x) = b_r x^r + \cdots + b_1 x + b_0 \qquad \text{and} \qquad h(x) = c_s x^s + \cdots + c_1 x + c_0.$$

Since $a_0 = b_0c_0$, $p \mid a_0$, and $p^2 \nmid a_0$, exactly one of b_0 and c_0 is divisible by p. Suppose, without loss of generality, that $p \mid b_0$ and $p \nmid c_0$.

Also $p \nmid a_n = b_r c_s$, so $p \nmid b_r$. Let t be the least index such that $p \nmid b_t$. Since $p \mid b_i$ for every $i < t$, the coefficient

$$a_t = b_t c_0 + b_{t-1}c_1 + \cdots + b_0 c_t$$

satisfies

$$p \mid a_t \qquad \text{and} \qquad p \mid b_{t-1}c_1 + \cdots + b_0 c_t.$$

Therefore $p \mid b_t c_0$. Since $p \nmid c_0$, this forces $p \mid b_t$, contradicting the choice of t. Hence $f(x)$ is irreducible over $\mathbb{Q}$. □

Example 20.3 Determine whether the following polynomials are irreducible over $\mathbb{Q}$.

1. $$f(x) = 5x^7 - 12x^6 + 6x^2 + 15x + 6.$$
2. $$f(x) = \frac{5}{2}x^5 + \frac{9}{2}x^4 + 15x^3 + \frac{9}{7}x^2 + 6x + \frac{3}{14}.$$

Solution. For (1), since $3 \nmid 5$, the prime 3 divides every other coefficient, and $3^2 \nmid 6$, Eisenstein's criterion with $p = 3$ shows that $f(x)$ is irreducible over $\mathbb{Q}$.

For (2), multiplying by 14 gives

$$h(x) = 14f(x) = 35x^5 + 63x^4 + 210x^3 + 18x^2 + 84x + 3.$$

Since $3 \nmid 35$, the prime 3 divides every other coefficient, and $3^2 \nmid 3$, Eisenstein's criterion shows that $h(x)$ is irreducible over $\mathbb{Q}$. Multiplication by a nonzero rational number does not affect irreducibility over $\mathbb{Q}$, so $f(x)$ is irreducible over $\mathbb{Q}$. □

Question 20.4 How does irreducibility of $f(x) \in F[x]$ affect the quotient ring $F[x]/\langle f(x)\rangle$?

Example 20.5 The quotient $\mathbb{R}[x]/\langle x^2-2\rangle$ is not an integral domain. Indeed, x^2-2 is reducible over $\mathbb{R}$, and in the quotient we have

$$\overline{x-\sqrt{2}}\,\overline{x+\sqrt{2}} = \overline{x^2-2} = \overline{0},$$

while neither factor is zero.

Theorem 20.6 Let F be a field and let $f(x) \in F[x]$. Then $\langle f(x)\rangle$ is maximal if and only if $f(x)$ is irreducible.

Proof. *Forward implication.* Suppose that $\langle f(x)\rangle$ is maximal. Then $f(x) \neq 0$ and $f(x)$ is not a unit, since $\langle 0\rangle$ is not maximal and $\langle f(x)\rangle = F[x]$ when $f(x)$ is a unit. Suppose

$$f(x) = g(x)h(x)$$

with $g(x), h(x) \in F[x]$. If $g(x)$ is not a unit, then

$$\langle f(x)\rangle \subseteq \langle g(x)\rangle \neq F[x].$$

By maximality, $\langle f(x)\rangle = \langle g(x)\rangle$. Hence $g(x) = f(x)u(x)$ for some $u(x) \in F[x]$. Comparing degrees in the equation $f(x) = g(x)h(x)$ gives $\deg(h) = 0$, so $h(x)$ is a unit. Thus $f(x)$ is irreducible.

Lecture 21 — Monday, February 29, 2016

Proof continued. *Reverse implication.* Suppose that $f(x)$ is irreducible. Let $I \trianglelefteq F[x]$ satisfy

$$\langle f(x)\rangle \subseteq I \subseteq F[x].$$

Since every ideal of $F[x]$ is principal, write $I = \langle g(x)\rangle$. Because $f(x) \in I$, there exists $h(x) \in F[x]$ such that

$$f(x) = g(x)h(x).$$

Since $f(x)$ is irreducible, either $g(x)$ is a unit or $h(x)$ is a unit. If $g(x)$ is a unit, then $I = F[x]$. If $h(x)$ is a unit, then $\langle f(x)\rangle = \langle g(x)\rangle = I$. Therefore $\langle f(x)\rangle$ is maximal. □

Corollary 21.1 Let F be a field and let $f(x) \in F[x]$. Then $F[x]/\langle f(x)\rangle$ is a field if and only if $f(x)$ is irreducible.

Proof. By Theorem 20.6, $f(x)$ is irreducible if and only if $\langle f(x)\rangle$ is maximal. By Theorem 10.8, this is equivalent to $F[x]/\langle f(x)\rangle$ being a field. □

Corollary 21.2 Let F be a field and let $p(x) \in F[x]$ be irreducible. If $a(x), b(x) \in F[x]$ and $p(x) \mid a(x)b(x)$, then $p(x) \mid a(x)$ or $p(x) \mid b(x)$.

Proof. Let $R = F[x]/\langle p(x)\rangle$. By Corollary 21.1, R is a field. Suppose that $p(x) \mid a(x)b(x)$. Then $a(x)b(x) \in \langle p(x)\rangle$, so in R we have

$$\overline{a(x)}\,\overline{b(x)} = \overline{0}.$$

Since R is a field, it is an integral domain. Hence $\overline{a(x)} = \overline{0}$ or $\overline{b(x)} = \overline{0}$. Therefore $a(x) \in \langle p(x)\rangle$ or $b(x) \in \langle p(x)\rangle$, which means $p(x) \mid a(x)$ or $p(x) \mid b(x)$. □

The punchline is that for a field F, the irreducible polynomials in $F[x]$ behave like prime numbers in $\mathbb{Z}$.

Example 21.3 The quotient $\mathbb{Z}[x]/\langle x+1\rangle$ is isomorphic to $\mathbb{Z}$, which is not a field. Thus $\langle x+1\rangle$ is not maximal in $\mathbb{Z}[x]$.

We use this theory to construct fields of size p^k, where p is prime and $k \in \mathbb{N}$.

Example 21.4 We construct a field of size 8. A field of size 8 must have characteristic 2, so we work over $\mathbb{Z}_2$. The polynomial

$$f(x) = x^3 + x + 1$$

has no roots in $\mathbb{Z}_2$, so it is irreducible over $\mathbb{Z}_2$ by Proposition 18.8. Hence

$$R = \mathbb{Z}_2[x]/\langle x^3 + x + 1\rangle$$

is a field by Corollary 21.1.

Every element of R has a representative of degree at most 2. Indeed, if $g(x) \in \mathbb{Z}_2[x]$, then by the division algorithm,

$$g(x) = f(x)q(x) + r(x),$$

where $r(x) = 0$ or $\deg(r) \leq 2$. Thus $\overline{g(x)} = \overline{r(x)}$ in R. Therefore

$$R = \{a + bx + cx^2 + \langle f(x)\rangle : a, b, c \in \mathbb{Z}_2\},$$

so $|R| = 2^3 = 8$.

Example 21.5 Since $x^2 + 1$ is irreducible over $\mathbb{Z}_3$, the quotient

$$\mathbb{Z}_3[x]/\langle x^2 + 1\rangle = \{a + bx + \langle x^2 + 1\rangle : a, b \in \mathbb{Z}_3\}$$

is a field of size 9. This agrees with the earlier isomorphism

$$\mathbb{Z}_3[x]/\langle x^2 + 1\rangle \cong \mathbb{Z}_3[i].$$

Divisibility in Integral Domains

Many notions discussed in $F[x]$ have analogues in $\mathbb{Z}$, including irreducible elements, prime elements, principal ideals, and unique factorization. We now generalize these notions to abstract integral domains.

Definition 21.6 Let R be an integral domain and let $a, b \in R$. We say that a and b are **associates** if $a = ub$ for some unit $u \in \mathcal{U}(R)$.

Definition 21.7 Let R be an integral domain. An element $a \in R$ is **irreducible** if $a \neq 0$, $a \notin \mathcal{U}(R)$, and whenever $a = bc$ with $b, c \in R$, at least one of b and c is a unit.

Definition 21.8 Let R be an integral domain. An element $a \in R$ is **prime** if $a \neq 0$, $a \notin \mathcal{U}(R)$, and whenever $a \mid bc$ with $b, c \in R$, we have $a \mid b$ or $a \mid c$.

Recall that in a ring R, we write $a \mid b$ if $b = ac$ for some $c \in R$.

Remark 21.9 Let R be an integral domain.

1. In $F[x]$, where F is a field, irreducible elements and prime elements coincide.
2. An irreducible element $a \in R$ need not generate a maximal ideal. For example, $\mathbb{Z}[x]/\langle x \rangle \cong \mathbb{Z}$ is not a field, so $\langle x \rangle$ is not maximal in $\mathbb{Z}[x]$.

The takeaway is that there is something "special" about $F[x]$ and $\mathbb{Z}$, which is the following:

Definition 21.10 An integral domain R is a **principal ideal domain** (**PID**) if every ideal of R is principal.

Example 21.11 If F is a field, then $F[x]$ is a principal ideal domain. The ring $\mathbb{Z}$ is also a principal ideal domain, and every field is a principal ideal domain.

Lecture 22 — Wednesday, March 02, 2016

We continue with irreducible and prime elements in integral domains, and with principal ideal domains.

Theorem 22.1 Let R be an integral domain. If $a \in R$ is prime, then a is irreducible.

Proof. Suppose that a is prime. Then $a \neq 0$ and $a \notin \mathcal{U}(R)$. If $a = bc$, then $a \mid bc$. Since a is prime, $a \mid b$ or $a \mid c$. Suppose, without loss of generality, that $a \mid b$. Then $b = ax$ for some $x \in R$, and hence

$$a = bc = axc.$$

Since R is an integral domain and $a \neq 0$, cancellation gives $1 = xc$. Thus c is a unit. Therefore a is irreducible. □

Theorem 22.2 Let R be a principal ideal domain. An element of R is prime if and only if it is irreducible.

Proof. *Forward implication.* This is Theorem 22.1.

Reverse implication. Suppose that $a \in R$ is irreducible. Let $a \mid bc$, where $b, c \in R$. Consider the ideal

$$I = \langle a, b \rangle.$$

Since R is a principal ideal domain, $I = \langle d \rangle$ for some $d \in R$. Since $a \in I$, there exists $r \in R$ such that $a = dr$. Since a is irreducible, either d is a unit or r is a unit.

If d is a unit, then $I = R$. Hence there exist $x, y \in R$ such that

$$ax + by = 1.$$

Multiplying by c gives

$$acx + bcy = c.$$

Since $a \mid acx$ and $a \mid bc$, it follows that $a \mid c$.

If r is a unit, then a and d are associates, so $\langle a \rangle = \langle d \rangle = I$. Since $b \in I$, we have $b \in \langle a \rangle$, and hence $a \mid b$.

Thus $a \mid b$ or $a \mid c$, so a is prime. □

Corollary 22.3 Let F be a field. Then any nonzero prime ideal in $F[x]$ is maximal.

Proof. Let P be a nonzero prime ideal of $F[x]$. Since $F[x]$ is a principal ideal domain, $P = \langle p(x) \rangle$ for some nonzero $p(x)$. The fact that P is proper makes $p(x)$ a nonunit, while primality of P makes $p(x)$ a prime element. By Theorem 22.2, $p(x)$ is irreducible. Hence P is maximal by Theorem 20.6. □

Theorem 22.4 Let R be a principal ideal domain and let $a \in R$ be nonzero. Then a is irreducible if and only if $\langle a \rangle$ is maximal.

Proof. *Forward implication.* Suppose that a is irreducible. Let $I \trianglelefteq R$ satisfy

$$\langle a \rangle \subseteq I \subseteq R.$$

Since R is a principal ideal domain, $I = \langle b \rangle$ for some $b \in R$. Because $a \in \langle b \rangle$, there exists $c \in R$ such that $a = bc$. Since a is irreducible, either b is a unit or c is a unit. If b is a unit, then $I = R$. If c is a unit, then a and b are associates, so $I = \langle b \rangle = \langle a \rangle$. Therefore $\langle a \rangle$ is maximal.

Reverse implication. Suppose that $\langle a \rangle$ is maximal. By Corollary 10.9, $\langle a \rangle$ is prime. Hence a is a prime element, and Theorem 22.1 shows that a is irreducible. □

Corollary 22.5 If R is a principal ideal domain, then every nonzero prime ideal is maximal.

Proof. Let P be a nonzero prime ideal. Since R is a principal ideal domain, $P = \langle a \rangle$ for some nonzero $a \in R$. The fact that P is proper makes a a nonunit, while primality of P makes a a prime element. Thus a is irreducible by Theorem 22.1, and $P = \langle a \rangle$ is maximal by Theorem 22.4. □

We now produce an example of an irreducible element that is not prime. Let $d \in \mathbb{Z}$ be squarefree with $d \neq 1$, and consider

$$\mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} : a, b \in \mathbb{Z}\}.$$

Define

$$N : \mathbb{Z}[\sqrt{d}] \to \mathbb{Z}_{\geq 0}, \quad N(a + b\sqrt{d}) = |a^2 - db^2|.$$

This function is the **norm** on $\mathbb{Z}[\sqrt{d}]$.

Proposition 22.6 Let $R = \mathbb{Z}[\sqrt{d}]$, where d is squarefree and $d \neq 1$. Then, for $x, y \in R$, the following hold:

1. $N(x) = 0$ if and only if $x = 0$.
2. $N(xy) = N(x)N(y)$.
3. $N(x) = 1$ if and only if $x \in \mathcal{U}(R)$.
4. If $N(x)$ is prime, then x is irreducible.

Proof. For (1), if $x = a + b\sqrt{d}$, then $N(x) = 0$ if and only if $a^2 = db^2$. Since d is squarefree, this forces $a = b = 0$. The converse is immediate.

For (2), write $x = a + b\sqrt{d}$ and $y = c + e\sqrt{d}$. Then

$$xy = (ac + bde) + (ae + bc)\sqrt{d}.$$

A direct calculation gives

$$(ac + bde)^2 - d(ae + bc)^2 = (a^2 - db^2)(c^2 - de^2),$$

and taking absolute values gives $N(xy) = N(x)N(y)$.

For (3), suppose that x is a unit, and choose $y \in R$ with $xy = 1$. Since $N(1) = 1$, (2) gives

$$1 = N(1) = N(xy) = N(x)N(y),$$

so $N(x) = 1$. Conversely, suppose that $N(x) = 1$ and write $x = a + b\sqrt{d}$. Let $\overline{x} = a - b\sqrt{d}$. Then

$$x\overline{x} = a^2 - db^2 = \pm 1,$$

so x is invertible in R.

For (4), suppose that $N(x)$ is prime and $x = ab$ with $a, b \in R$. Then

$$N(x) = N(a)N(b).$$

Since $N(x)$ is prime in $\mathbb{Z}$, one of $N(a)$ and $N(b)$ is equal to 1. By (3), one of a and b is a unit. Thus x is irreducible. □

Remark 22.7 If a and b are associates in $\mathbb{Z}[\sqrt{d}]$, then $N(a) = N(b)$. The converse is false in general.

Example 22.8 The element $1 + \sqrt{-3}$ is irreducible but not prime in $\mathbb{Z}[\sqrt{-3}]$. Its norm is

$$N(1 + \sqrt{-3}) = 4,$$

so it is not a unit. Suppose that $1 + \sqrt{-3} = xy$ with $x, y \in \mathbb{Z}[\sqrt{-3}]$. Then $4 = N(x)N(y)$. There is no element of norm 2, because

$$N(a + b\sqrt{-3}) = 2$$

would imply $a^2 + 3b^2 = 2$, which has no integer solutions. Therefore one of $N(x)$ and $N(y)$ is equal to 1, so one of x and y is a unit. Thus $1 + \sqrt{-3}$ is irreducible.

However,

$$(1 + \sqrt{-3})(1 - \sqrt{-3}) = 4 = 2 \cdot 2,$$

so $(1 + \sqrt{-3}) \mid 2 \cdot 2$. We claim that $(1 + \sqrt{-3}) \nmid 2$. If it did, then

$$2 = (1 + \sqrt{-3})(a + b\sqrt{-3})$$

for some $a, b \in \mathbb{Z}$. Comparing coefficients gives

$$2 = a - 3b \qquad \text{and} \qquad 0 = a + b.$$

Thus $a = 1/2$, which is impossible. Hence $(1 + \sqrt{-3}) \nmid 2$, and $1 + \sqrt{-3}$ is not prime.

Lecture 23 — Friday, March 04, 2016

Unique Factorization Domains

Example 23.1 The element 7 is irreducible in $\mathbb{Z}[\sqrt{5}]$. Suppose that $7 = xy$. Then

$$49 = N(7) = N(x)N(y).$$

It is impossible for an element to have norm 7: if $x = a + b\sqrt{5}$ and $N(x) = 7$, then

$$a^2 - 5b^2 = \pm 7.$$

Reducing modulo 5 gives $a^2 \equiv \pm 2 \pmod 5$, but 2 and 3 are not quadratic residues modulo 5. Hence the only possible norm factors are 1 and 49. Therefore one of x and y has norm 1, so one of them is a unit. Thus 7 is irreducible.

Example 23.2 It is not always true that $\mathcal{U}(\mathbb{Z}[\sqrt{d}]) = \{1, -1\}$. In $\mathbb{Z}[\sqrt{5}]$,

$$N(2 + \sqrt{5}) = |4 - 5| = 1,$$

so $2 + \sqrt{5}$ is a unit.

If F is a field, then $F[x]$ is a principal ideal domain, so irreducible elements are prime in $F[x]$. We now ask whether $F[x]$ has unique factorization.

Definition 23.3 An integral domain R is a **unique factorization domain** (**UFD**) if the following hold:

1. Every nonzero nonunit element of R can be written as a product of irreducibles; and
2. Such factorizations into irreducibles are unique up to reordering and replacing factors by associates.

Example 23.4 The ring $\mathbb{Z}$ is a unique factorization domain, because every nonzero nonunit integer can be written as a product of prime numbers, and this factorization is unique up to reordering and replacing factors by ± 1.

Example 23.5 The ring $\mathbb{Z}[x]$ is a unique factorization domain. We will skip the proof here, which is essentially an involved application of Gauss's lemma.

Example 23.6 Every field is a unique factorization domain, vacuously, because it has no nonzero nonunit elements.

Definition 23.7 Let R be a ring. We say that R satisfies the ascending chain condition if whenever

$$I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$$

where $I_i \trianglelefteq R$ for each i, there exists N such that $I_N = I_{N+1} = I_{N+2} = \cdots$. Similarly, we say that R satisfies the descending chain condition if whenever

$$I_1 \supseteq I_2 \supseteq I_3 \supseteq \cdots$$

where $I_i \trianglelefteq R$ for each i, there exists N such that $I_N = I_{N+1} = I_{N+2} = \cdots$. A ring is **Noetherian** if it satisfies the ascending chain condition, and is **Artinian** if it satisfies the descending chain condition.

Proposition 23.8 Any field is both Noetherian and Artinian.

Proof. A field has only two ideals, $\{0\}$ and the field itself. Thus any ascending or descending chain of ideals must stabilize immediately. □

Lemma 23.9 Every principal ideal domain is Noetherian.

Proof. Let R be a principal ideal domain, and let

$$I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$$

be an ascending chain of ideals. Set

$$I = \bigcup_{i=1}^{\infty} I_i.$$

Then I is an ideal. Since R is a principal ideal domain, $I = \langle a \rangle$ for some $a \in R$. The element a belongs to I_N for some N, and hence

$$I = \langle a \rangle \subseteq I_N.$$

Thus $I_N = I_{N+1} = I_{N+2} = \cdots$, so every ascending chain of ideals stabilizes. □

Theorem 23.10 Every principal ideal domain is a unique factorization domain.

Proof. Let $a_0 \in R$ be a nonzero nonunit. We first show that a_0 has an irreducible factor. If a_0 is irreducible, there is nothing to prove. Otherwise, write $a_0 = b_1 a_1$, where neither b_1 nor a_1 is a unit. If a_1 is not irreducible, write $a_1 = b_2 a_2$ with b_2 and a_2 nonunits. Continuing in this way would produce a strictly ascending chain

$$\langle a_0 \rangle \subsetneq \langle a_1 \rangle \subsetneq \langle a_2 \rangle \subsetneq \cdots .$$

This contradicts Lemma 23.9. Hence the process stops, and some factor is irreducible.

Now write $a_0 = p_1 c_1$, where p_1 is irreducible. If c_1 is a unit, then a_0 is already a product of irreducibles. Otherwise, apply the previous paragraph to c_1 and continue. The same argument using an ascending chain shows that this process must stop, so a_0 is a product of irreducibles.

For uniqueness, suppose

$$a = p_1 p_2 \cdots p_t = q_1 q_2 \cdots q_s$$

are two factorizations into irreducibles. By Theorem 22.2, each p_i and q_j is prime. Since $p_1 \mid q_1 q_2 \cdots q_s$, we have $p_1 \mid q_j$ for some j. After reordering, assume $p_1 \mid q_1$. Since q_1 is irreducible, p_1 and q_1 are associates. Cancelling associate factors reduces the equality to a shorter equality of products of irreducibles. Repeating this argument shows that the two factorizations agree up to reordering and associates. □

Example 23.11 If F is a field, then $F[x]$ is a principal ideal domain, and hence $F[x]$ is a unique factorization domain.

Example 23.12 The ring $\mathbb{Z}[x]$ is a unique factorization domain but not a principal ideal domain.

Example 23.13 The ring $\mathbb{Z}[\sqrt{-5}]$ is an integral domain but not a unique factorization domain. Indeed,

$$(1+\sqrt{-5})(1-\sqrt{-5}) = 6 = 2 \cdot 3.$$

We claim that $1+\sqrt{-5}$, $1-\sqrt{-5}$, 2, and 3 are irreducible. Their norms are

$$N(1+\sqrt{-5}) = N(1-\sqrt{-5}) = 6, \qquad N(2) = 4, \qquad \text{and} \qquad N(3) = 9.$$

If one of these elements factored nontrivially, then the norms of the two nonunit factors would have to include a factor of norm 2 or 3. But

$$N(a+b\sqrt{-5}) = a^2 + 5b^2$$

is never equal to 2 or 3 for integers a, b. Hence all four elements are irreducible. The displayed equality therefore gives two distinct factorizations of 6 into irreducibles.

Lecture 24 — Monday, March 07, 2016

Euclidean Domains

Remark 24.1 Let $R = \mathbb{Z}[\sqrt{d}]$, where d is squarefree. If $a, b \in R$ are associates, then $N(a) = N(b)$. Therefore $N(a) \neq N(b)$ implies that a and b are not associates. The converse need not hold.

Our next generalization is to integral domains that admit a division algorithm.

Definition 24.2 An integral domain R is a **Euclidean domain** if there exists a function

$$d : R \setminus \{0\} \to \mathbb{N} \cup \{0\}$$

such that the following hold:

1. $d(a) \leq d(ab)$ for all $a, b \in R \setminus \{0\}$.
2. For all $a, b \in R$ with $b \neq 0$, there exist $q, r \in R$ such that

$$a = bq + r,$$

where either $r = 0$ or $d(r) < d(b)$.

The function d is called a **measure**.

Note that the elements q and r in Definition 24.2 are *not* required to be unique.

Remark 24.3 The ring $\mathbb{Z}$ is a Euclidean domain with measure $d(a) = |a|$. Indeed, if $a, b \in \mathbb{Z}$ and $b \neq 0$, then there exist unique $q, r \in \mathbb{Z}$ such that

$$a = bq + r, \quad 0 \leq r < |b|.$$

Example 24.4 If F is a field, then $F[x]$ is a Euclidean domain with measure $d(f(x)) = \deg(f(x))$.

Example 24.5 The Gaussian integers $\mathbb{Z}[i]$ form a Euclidean domain with measure

$$d(a + bi) = a^2 + b^2 = N(a + bi).$$

For nonzero $a, b \in \mathbb{Z}[i]$, we have

$$d(ab) = N(ab) = N(a)N(b) \geq N(a) = d(a).$$

Now let $x, y \in \mathbb{Z}[i]$ with $y \neq 0$. In $\mathbb{Q}[i]$, write

$$xy^{-1} = s + ti$$

with $s, t \in \mathbb{Q}$. Choose integers n and m nearest to s and t, respectively. Then

$$xy^{-1} = (n + mi) + (s - n) + (t - m)i,$$

so

$$x = (n + mi)y + \big((s - n) + (t - m)i\big)y.$$

Let

$$q = n + mi \qquad \text{and} \qquad r = \big((s - n) + (t - m)i\big)y.$$

Then $x = qy + r$. If $r \neq 0$, then

$$\begin{aligned} d(r) &= N\big((s - n) + (t - m)i\big)N(y) \\ &= \big((s - n)^2 + (t - m)^2\big)N(y) \\ &\leq \left(\frac{1}{4} + \frac{1}{4}\right) N(y) = \frac{1}{2}d(y) < d(y). \end{aligned}$$

Thus $\mathbb{Z}[i]$ is a Euclidean domain.

Theorem 24.6 Every Euclidean domain is a principal ideal domain.

Proof. Let R be a Euclidean domain with measure d, and let $I \trianglelefteq R$ be a nonzero ideal. Choose $a \in I \setminus \{0\}$ with $d(a)$ minimal among the measures of nonzero elements of I.

Now take $b \in I$. Because $a \in I$ and I is an ideal, we have $\langle a \rangle \subseteq I$. By the Euclidean property, there exist $q, r \in R$ such that

$$b = aq + r,$$

where $r = 0$ or $d(r) < d(a)$. Since $r = b - aq \in I$, the minimality of $d(a)$ forces $r = 0$. Hence $b = aq \in \langle a \rangle$. Thus $I = \langle a \rangle$, so I is principal. □

Corollary 24.7 Every Euclidean domain is a unique factorization domain.

Proof. This follows from Theorem 24.6 and Theorem 23.10. □

Example 24.8 The ring $\mathbb{Z}[i]$ is a principal ideal domain, hence a unique factorization domain.

We have shown that every Euclidean domain is a principal ideal domain, and every principal ideal domain is a unique factorization domain. However, the converses are false. For example, $\mathbb{Z}[x]$ is a unique factorization domain but not a principal ideal domain.

The hierarchy in Figure 9 keeps the valid implications separate from their failed converses.

FIGURE 9

Example 24.9 Let

$$\theta = \frac{1+\sqrt{-19}}{2}.$$

Then $\mathbb{Z}[\theta]$ is a principal ideal domain, but it is not a Euclidean domain.

4. Field Extensions

Field Extensions and Splitting Fields

Definition 24.10 A field K is an **extension field**, or **field extension**, of a field F if $F \subseteq K$ and the operations on F are the restrictions of the operations on K. Equivalently, F is a subfield of K. In this situation, we write $F \subseteq K$.

Example 24.11 We have $\mathbb{R} \subseteq \mathbb{C}$, $\mathbb{Q} \subseteq \mathbb{Q}(x)$, and $\mathbb{Z}_3 \subseteq \mathbb{Z}_3(x)$.

Let F be a field and let $f(x) \in F[x]$ be irreducible. Then $F[x]/\langle f(x)\rangle$ is a field by Corollary 21.1. We view it as a field extension of F through the natural embedding

$$F \to F[x]/\langle f(x)\rangle, \quad a \mapsto \overline{a}.$$

Lecture 25 — Wednesday, March 09, 2016

Notation 25.1 The notation $F \subseteq K$ may mean either that F is literally a subfield of K, or that F is identified with a naturally embedded isomorphic copy inside K.

Example 25.2 The containment $\mathbb{Q} \subseteq \mathbb{C}$ is literal. The containment $\mathbb{Z}_5 \subseteq \mathbb{Z}_5(x)$ is via $[a]_5 \mapsto [a]_5$. The containment

$$\mathbb{Q} \subseteq \mathbb{Q}[x]/\langle x^2 - 2\rangle$$

is via $a \mapsto \overline{a}$.

Remark 25.3 If F is a subfield of K, then every polynomial in $F[x]$ may be viewed as a polynomial in $K[x]$. If instead F is identified with a subfield of K through an isomorphism $\varphi : F \to \varphi(F) \subseteq K$, then

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 \in F[x]$$

corresponds to

$$\widetilde{f}(x) = \sum_{i=0}^{n} \varphi(a_i) x^i \in K[x].$$

There are two main motivations for studying field extensions.

1. We can find roots of polynomials in larger fields.
2. If $F \subseteq K$, then K is an F-vector space, where vector addition is addition in K and scalar multiplication is multiplication in K.

Theorem 25.4 (Kronecker's theorem) Let F be a field and let $f(x) \in F[x]$ be nonconstant. Then there exists an extension field E of F such that $f(x)$ has a root in E.

Proof. Since $F[x]$ is a unique factorization domain, $f(x)$ has an irreducible factor $p(x)$. Write $f(x) = p(x)h(x)$ in $F[x]$. Consider

$$E = F[y]/\langle p(y)\rangle.$$

Since $p(y)$ is irreducible over F, Corollary 21.1 shows that E is a field. Also F embeds into E by $a \mapsto \overline{a}$.

The element $\overline{y} \in E$ is a root of $p(x)$, because

$$p(\overline{y}) = \overline{p(y)} = \overline{0}.$$

Therefore

$$f(\overline{y}) = p(\overline{y})h(\overline{y}) = 0,$$

so $f(x)$ has a root in the extension field E. □

Corollary 25.5 Let R be an integral domain and let $f(x) \in R[x]$ be nonconstant. There exists a field K with $R \subseteq K$ such that $f(x)$ has a root in K.

Proof. Let Q be the field of fractions of R. Then $R \subseteq Q$ and $f(x) \in Q[x]$. Apply Theorem 25.4. □

Example 25.6 The integral domain hypothesis in Corollary 25.5 is necessary. Consider $f(x) = 2x + 1 \in \mathbb{Z}_4[x]$. If $f(x)$ had a root a in a ring extension of $\mathbb{Z}_4$, then $2a + 1 = 0$. Multiplying by 2 gives $2 = 0$, a contradiction to the copy of $\mathbb{Z}_4$ inside the extension.

Definition 25.7 Let $F \subseteq K$ be a field extension and let $\alpha \in K$. We denote by $F(\alpha)$ the smallest subfield of K containing both F and α, and we say that $F(\alpha)$ is obtained by **adjoining** α to F.

This field exists because it is the intersection of all subfields of K that contain F and α:

$$F(\alpha) = \bigcap \{E : F \subseteq E \subseteq K,\ E \text{ is a field},\ \alpha \in E\}.$$

Example 25.8 We have

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}.$$

If $a + b\sqrt{2} \neq 0$, then

$$\frac{1}{a+b\sqrt{2}} = \frac{a-b\sqrt{2}}{a^2-2b^2},$$

so this set is closed under inverses.

Example 25.9 The field $\mathbb{Q}(x)$ is the field of fractions of $\mathbb{Q}[x]$.

Example 25.10 $\mathbb{R}(i) = \mathbb{C}$.

Definition 25.11 Let K be a field extension of F, and let $\alpha_1, \ldots, \alpha_n \in K$. Then $F(\alpha_1, \ldots, \alpha_n)$ is the smallest subfield of K containing F and all of $\alpha_1, \ldots, \alpha_n$.

Example 25.12 The smallest field extension of $\mathbb{Q}$ in which $x^2 - 2$ has a root is $\mathbb{Q}(\sqrt{2})$.

Definition 25.13 Let $F \subseteq K$ be a field extension, and let $f(x) \in F[x]$ be nonconstant. We say that $f(x)$ **splits** over K if it factors over K into linear factors. The field K is a **splitting field** for $f(x)$ over F if $f(x)$ splits over K and K is generated over F by the roots of $f(x)$.

Example 25.14 The splitting field of $(x^2-2)(x^2-3)$ over $\mathbb{Q}$ is $\mathbb{Q}(\sqrt{2}, \sqrt{3})$. The splitting field of $x^2 + 1$ over $\mathbb{Q}$ is $\mathbb{Q}(i)$.

In general, if $f(x)$ splits over K, then

$$f(x) = b(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)$$

for roots $\alpha_1, \ldots, \alpha_n \in K$. The splitting field is therefore $F(\alpha_1, \ldots, \alpha_n)$.

Theorem 25.15 (Existence of splitting fields) Let F be a field and let $f(x) \in F[x]$ be nonconstant. Then there exists a splitting field for $f(x)$ over F.

Lecture 26 — Friday, March 11, 2016

Proof. We proceed by induction on $\deg(f)$. If $\deg(f) = 1$, then F is a splitting field for $f(x)$ over F.

Suppose now that $\deg(f) > 1$ and that splitting fields exist for all polynomials of smaller degree over any field. By Theorem 25.4, there is an extension field E of F in which $f(x)$ has a root α. Then

$$f(x) = (x - \alpha)g(x)$$

in $E[x]$, where $\deg(g) < \deg(f)$. By the induction hypothesis, $g(x)$ has a splitting field K over E. If the roots of $g(x)$ in K are $b_1, \ldots, b_m$, then

$$F(\alpha, b_1, \ldots, b_m)$$

is a splitting field for $f(x)$ over F. □

Example 26.1 Find the splitting field for $x^4 + x^2 + 1$ over $\mathbb{Q}$. We factor

$$x^4 + x^2 + 1 = (x^2 + x + 1)(x^2 - x + 1).$$

The roots of $x^2 + x + 1$ are $\dfrac{-1 \pm \sqrt{-3}}{2}$, and the roots of $x^2 - x + 1$ are $\dfrac{1 \pm \sqrt{-3}}{2}$. Thus the splitting field is

$$\mathbb{Q}(\sqrt{-3}) = \mathbb{Q}(i\sqrt{3}).$$

Example 26.2 We have

$$\mathbb{Q}(2\sqrt{2}, \sqrt{2}) = \mathbb{Q}(\sqrt{2}).$$

Indeed, $\sqrt{2} \in \mathbb{Q}(2\sqrt{2}, \sqrt{2})$, so $\mathbb{Q}(\sqrt{2}) \subseteq \mathbb{Q}(2\sqrt{2}, \sqrt{2})$. Conversely, both $\sqrt{2}$ and $2\sqrt{2}$ lie in $\mathbb{Q}(\sqrt{2})$, so $\mathbb{Q}(2\sqrt{2}, \sqrt{2}) \subseteq \mathbb{Q}(\sqrt{2})$.

Example 26.3 We show that

$$\mathbb{Q}(i, \sqrt{3}) = \mathbb{Q}(i + \sqrt{3}).$$

The inclusion $\mathbb{Q}(i + \sqrt{3}) \subseteq \mathbb{Q}(i, \sqrt{3})$ is immediate. Conversely, since $i + \sqrt{3} \in \mathbb{Q}(i + \sqrt{3})$, we

have

$$(i+\sqrt{3})^2 = 2 + 2i\sqrt{3} \in \mathbb{Q}(i+\sqrt{3}),$$

and hence $i\sqrt{3} \in \mathbb{Q}(i+\sqrt{3})$. Therefore

$$(i\sqrt{3})(i+\sqrt{3}) = -\sqrt{3} + 3i \in \mathbb{Q}(i+\sqrt{3}).$$

Combining this with $i+\sqrt{3}$ gives

$$3(i+\sqrt{3}) - (-\sqrt{3}+3i) = 4\sqrt{3} \in \mathbb{Q}(i+\sqrt{3}).$$

Together with $i+\sqrt{3}$, this lets us solve for both i and $\sqrt{3}$, so $i, \sqrt{3} \in \mathbb{Q}(i+\sqrt{3})$. Hence $\mathbb{Q}(i,\sqrt{3}) \subseteq \mathbb{Q}(i+\sqrt{3})$.

Theorem 26.4 Let F be a field and let $p(x) \in F[x]$ be irreducible. Suppose that α lies in some extension $F \subseteq K$ and that $p(\alpha) = 0$. Then

$$F(\alpha) \cong F[x]/\langle p(x)\rangle.$$

Moreover, if $\deg(p) = n$, then every element of $F(\alpha)$ can be written uniquely in the form

$$c_{n-1}\alpha^{n-1} + c_{n-2}\alpha^{n-2} + \cdots + c_1\alpha + c_0,$$

where $c_i \in F$.

Proof. Define $\ell : F[x] \to F(\alpha)$ by $\ell(f(x)) = f(\alpha)$. This is a homomorphism. Since $p(\alpha) = 0$, we have $\langle p(x)\rangle \subseteq \ker \ell$. The ideal $\langle p(x)\rangle$ is maximal by Theorem 20.6, and $\ker \ell \neq F[x]$ because $\ell(1) = 1$. Therefore

$$\ker \ell = \langle p(x)\rangle.$$

By the first isomorphism theorem, Theorem 14.1,

$$F[x]/\langle p(x)\rangle \cong \ell(F[x]).$$

The image $\ell(F[x])$ contains F and α, because $\ell(a) = a$ for $a \in F$ and $\ell(x) = \alpha$. Since $F[x]/\langle p(x)\rangle$ is a field, $\ell(F[x])$ is a field. By minimality of $F(\alpha)$, we get $\ell(F[x]) = F(\alpha)$.

The isomorphism is explicitly

$$\Psi : F[x]/\langle p(x)\rangle \to F(\alpha), \quad \Psi(\overline{f(x)}) = f(\alpha).$$

By the division algorithm, each element in $F[x]/\langle p(x)\rangle$ can be written as

$$\overline{c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \cdots + c_1 x + c_0},$$

where $c_i \in F$. Since Ψ is an isomorphism, each element of $F(\alpha)$ can be written as

$$\Psi(\overline{c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \cdots + c_1 x + c_0}) = c_{n-1}\alpha^{n-1} + c_{n-2}\alpha^{n-2} + \cdots + c_1\alpha + c_0.$$

If there were two distinct expressions for the same element of $F(\alpha)$, then their difference would give a nontrivial linear relation among $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$, which would contradict the irreducibility of $p(x)$. Thus the expression is unique. □

Corollary 26.5 Let F be a field and let $p(x) \in F[x]$ be irreducible. If α and β lie in extensions of F and satisfy $p(\alpha) = p(\beta) = 0$, then $F(\alpha) \cong F(\beta)$.

Proof. By Theorem 26.4,

$$F(\alpha) \cong F[x]/\langle p(x)\rangle \cong F(\beta).$$

□

Example 26.6 Let $\alpha = \sqrt[6]{2} = 2^{1/6}$. Then α is a root of $x^6 - 2$, which is irreducible over $\mathbb{Q}$ by Eisenstein's criterion with $p = 2$. By Theorem 26.4,

$$\mathbb{Q}(\sqrt[6]{2}) \cong \mathbb{Q}[x]/\langle x^6 - 2\rangle.$$

Also,

$$\mathbb{Q}(\sqrt[6]{2}) = \{c_5 2^{5/6} + c_4 2^{4/6} + c_3 2^{3/6} + c_2 2^{2/6} + c_1 2^{1/6} + c_0 : c_i \in \mathbb{Q}\}.$$

Example 26.7 Since $x^2 - 2$ is irreducible over $\mathbb{Q}$,

$$\mathbb{Q}(\sqrt{2}) \cong \mathbb{Q}[x]/\langle x^2 - 2\rangle,$$

and every element of $\mathbb{Q}(\sqrt{2})$ has the form $c_1\sqrt{2} + c_0$ with $c_0, c_1 \in \mathbb{Q}$.

Example 26.8 Let $\alpha = \sqrt{1+\sqrt{5}}$. Then

$$\alpha^2 = 1 + \sqrt{5},$$

so

$$(\alpha^2 - 1)^2 = 5.$$

Thus

$$\alpha^4 - 2\alpha^2 - 4 = 0.$$

Taking $p(x) = x^4 - 2x^2 - 4$, we have $p(\alpha) = 0$. The rational-root candidates $\pm 1, \pm 2$, and ± 4 are not roots of $p(x)$, so any factorization over $\mathbb{Q}$ would have two monic quadratic factors. Since the coefficients of x^3 and x vanish, such a factorization would have the form

$$p(x) = (x^2 + ux + v)(x^2 - ux + w),$$

where the coefficient of x gives $u(w - v) = 0$. If $u = 0$, then $v + w = -2$ and $vw = -4$, so v and w would be the irrational roots $-1 \pm \sqrt{5}$ of $t^2 + 2t - 4$. If $v = w$, then $v^2 = -4$, which is impossible over $\mathbb{Q}$. Therefore $p(x)$ is irreducible over $\mathbb{Q}$.

It now follows from Theorem 26.4 that

$$\mathbb{Q}(\sqrt{1+\sqrt{5}}) \cong \mathbb{Q}[x]/\langle x^4 - 2x^2 - 4\rangle,$$

and every element has the form

$$c_3\alpha^3 + c_2\alpha^2 + c_1\alpha + c_0,$$

where $c_i \in \mathbb{Q}$.

Lecture 27 — Monday, March 14, 2016

Uniqueness of Splitting Fields

Example 27.1 We find the splitting field of x^4+1 over $\mathbb{Q}$. The roots of x^4+1 in $\mathbb{C}$ are the fourth roots of

$$-1 = \cos(\pi) + i\sin(\pi).$$

They are

$$\cos\left(\frac{\pi+2k\pi}{4}\right) + i\sin\left(\frac{\pi+2k\pi}{4}\right), \quad k=0,1,2,3.$$

Thus the roots are

$$\frac{1}{\sqrt{2}} \pm \frac{i}{\sqrt{2}}, \qquad \text{and} \qquad -\frac{1}{\sqrt{2}} \pm \frac{i}{\sqrt{2}}.$$

Since $1/\sqrt{2} = \sqrt{2}/2$, the splitting field is $\mathbb{Q}(i,\sqrt{2})$.

If F is a field, $p(x) \in F[x]$ is irreducible of degree n, and $p(\alpha)=0$ in some extension of F, then

$$F(\alpha) \cong F[x]/\langle p(x)\rangle$$

and every element of $F(\alpha)$ has a unique expression

$$c_0 + c_1\alpha + \cdots + c_{n-1}\alpha^{n-1}, \quad c_i \in F.$$

Thus $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ is a basis for $F(\alpha)$ over F.

We now prove that splitting fields are unique up to isomorphism.

Lemma 27.2 Let F and F' be fields, let $\varphi : F \to F'$ be an isomorphism, and extend φ to an isomorphism $F[x] \to F'[x]$ by applying φ to coefficients. Let $p(x) \in F[x]$ be irreducible. Suppose that a is a root of $p(x)$ in an extension of F, and that b is a root of $\varphi(p(x))$ in an extension of F'. Then there exists an isomorphism

$$\Psi : F(a) \to F'(b)$$

such that $\Psi|_F = \varphi$ and $\Psi(a) = b$.

Proof. The coefficientwise extension $\varphi : F[x] \to F'[x]$ is an isomorphism, and it sends irreducible

polynomials to irreducible polynomials. By Theorem 26.4, we have isomorphisms

$$\theta_a : F[x]/\langle p(x)\rangle \to F(a), \qquad \text{and} \qquad \theta_b : F'[x]/\langle \varphi(p(x))\rangle \to F'(b),$$

where $\theta_a(\overline{x}) = a$ and $\theta_b(\overline{x}) = b$.

Define

$$\overline{\varphi} : F[x]/\langle p(x)\rangle \to F'[x]/\langle \varphi(p(x))\rangle$$

by

$$\overline{\varphi}(\overline{f(x)}) = \overline{\varphi(f(x))}.$$

This is well-defined because $f(x) \in \langle p(x)\rangle$ if and only if $\varphi(f(x)) \in \langle \varphi(p(x))\rangle$. It is an isomorphism.

The isomorphisms are summarized in Figure 10.

$$F(a) \xrightarrow[\cong]{\theta_a^{-1}} \frac{F[x]}{\langle p(x)\rangle} \xrightarrow[\cong]{\overline{\varphi}} \frac{F'[x]}{\langle \varphi(p(x))\rangle} \xrightarrow[\cong]{\theta_b} F'(b)$$

$$a \longmapsto \overline{x} \longmapsto \overline{x} \longmapsto b$$

FIGURE 10

Set

$$\Psi = \theta_b \circ \overline{\varphi} \circ \theta_a^{-1}.$$

Then Ψ is an isomorphism by Proposition 12.4(4) because it is a composition of isomorphisms, $\Psi(a) = b$ because $\theta_b(\overline{x}) = b$, and $\Psi|_F = \varphi$ because $\overline{\varphi}(\overline{f(x)}) = \overline{\varphi(f(x))}$ for all $f(x) \in F[x]$. □

Theorem 27.3 Let F and F' be fields, let $\varphi : F \to F'$ be an isomorphism, and let $f(x) \in F[x]$ be nonconstant. If E is a splitting field for $f(x)$ over F and E' is a splitting field for $\varphi(f(x))$ over F', then there exists an isomorphism

$$\Psi : E \to E'$$

such that $\Psi|_F = \varphi$.

Proof. We proceed by induction on $\deg(f)$. If $\deg(f) = 1$, then $E = F$ and $E' = F'$, so we may take $\Psi = \varphi$.

Suppose that $\deg(f) > 1$ and that the result holds for polynomials of smaller degree. Let $p(x)$ be an irreducible factor of $f(x)$, and let $a \in E$ be a root of $p(x)$. Since E' is a splitting field for $\varphi(f(x))$, the polynomial $\varphi(p(x))$ has a root $b \in E'$. By Lemma 27.2, there is an isomorphism

$$\alpha : F(a) \to F'(b)$$

such that $\alpha|_F = \varphi$ and $\alpha(a) = b$.

In $F(a)[x]$, write

$$f(x) = (x - a)g(x).$$

Then E is a splitting field for $g(x)$ over $F(a)$, and E' is a splitting field for $\alpha(g(x))$ over $F'(b)$. By induction, there exists an isomorphism $\Psi : E \to E'$ such that $\Psi|_{F(a)} = \alpha$. Therefore $\Psi|_F = \varphi$. □

Corollary 27.4 Let F be a field and let $f(x) \in F[x]$ be nonconstant. Then any two splitting fields for $f(x)$ over F are isomorphic.

Proof. Apply Theorem 27.3 with $F = F'$ and $\varphi = \mathrm{id}_F$. □

Lecture 28 — Wednesday, March 16, 2016

Types of Extensions

Let F be a field and let $f(x) \in F[x]$ be irreducible of degree n. Our previous work shows that if $f(a) = 0$ in an extension of F, then

$$F(a) \cong F[x]/\langle f(x) \rangle,$$

and every element of $F(a)$ has a unique expression

$$c_0 + c_1 a + \cdots + c_{n-1} a^{n-1}, \quad c_i \in F.$$

There are two main motivations for studying extension fields: splitting fields and linear algebra over field extensions.

Definition 28.1 Let $F \subseteq K$ be a field extension, and let $a \in K$. We say that a is **algebraic** over F if there is a nonzero polynomial $f(x) \in F[x]$ such that $f(a) = 0$. Otherwise, we say that a is **transcendental** over F.

The extension $F \subseteq K$ is called an **algebraic extension** if every element of K is algebraic over F. Otherwise, it is called a **transcendental extension**.

Charles Hermite proved that e is transcendental over $\mathbb{Q}$ in 1873, and Ferdinand von Lindemann proved that π is transcendental over $\mathbb{Q}$ in 1882. Thus we can say that $\mathbb{Q} \subseteq \mathbb{Q}(\pi)$ is a transcendental extension.

Theorem 28.2 Let $F \subseteq K$ be a field extension, and let $a \in K$. If a is algebraic over F, then

$$F(a) \cong F[x]/\langle p(x)\rangle,$$

where $p(x)$ is a polynomial in $F[x]$ of minimal degree such that $p(a) = 0$. Moreover, $p(x)$ is irreducible.

Proof. Define $\ell : F[x] \to F(a)$ by $\ell(f(x)) = f(a)$. Since a is algebraic over F, the kernel of ℓ is nonzero. Thus $\ker(\ell)$ is a nonzero ideal of $F[x]$, so

$$\ker(\ell) = \langle p(x)\rangle$$

for some nonzero polynomial $p(x) \in \ker(\ell)$ of minimal degree. By the first isomorphism theorem,

$$F[x]/\langle p(x)\rangle \cong \ell(F[x]).$$

Also, $p(a) = 0$ because $p(x) \in \ker(\ell)$.

The polynomial $p(x)$ is irreducible. Indeed, if $p(x) = g(x)h(x)$ with both factors of positive degree, then

$$0 = p(a) = g(a)h(a).$$

Since K is a field, either $g(a) = 0$ or $h(a) = 0$, contradicting the minimality of $\deg(p)$. Therefore $F[x]/\langle p(x)\rangle$ is a field, and hence $\ell(F[x])$ is a field contained in $F(a)$. Since $\ell(F[x])$ contains both F and a, it contains the smallest field containing F and a. Thus $\ell(F[x]) = F(a)$, and the desired isomorphism follows. □

Corollary 28.3 If $a \in K$ is algebraic over F, then there is a unique monic irreducible polynomial $p(x) \in F[x]$ such that $p(a) = 0$.

Proof. In the proof of Theorem 28.2, the polynomials in $\ker(\ell)$ of minimal degree are precisely the generators of the principal ideal $\ker(\ell)$. Therefore any two of them are associates in $F[x]$. Exactly one associate is monic. □

Definition 28.4 The unique monic irreducible polynomial $p(x) \in F[x]$ such that $p(a) = 0$ is called the **minimal polynomial** for a over F.

Proposition 28.5 Let $a \in K$ be algebraic over F, and let $p(x)$ be the minimal polynomial for a over F. If $f(x) \in F[x]$ satisfies $f(a) = 0$, then $p(x) \mid f(x)$.

Proof. By the division algorithm in $F[x]$, there are $q(x), r(x) \in F[x]$ such that

$$f(x) = p(x)q(x) + r(x), \quad \deg(r) < \deg(p),$$

unless $r(x) = 0$. Evaluating at a gives $r(a) = 0$. If $r(x) \neq 0$, then this contradicts the minimality of the degree of $p(x)$. Hence $r(x) = 0$, and $p(x) \mid f(x)$. □

Remark 28.6 If $F \subseteq K$ is a field extension, then K is an F-vector space: the elements of K are the vectors, and the elements of F act as scalars.

Definition 28.7 Let $F \subseteq K$ be a field extension. We say that K has **degree n over F**, denoted $[K : F] = n$, if K has dimension n as an F-vector space. If $[K : F] < \infty$, then $F \subseteq K$ is called a **finite extension**; otherwise, it is called an **infinite extension**.

Example 28.8 If $a \in K$ is algebraic over F, and if $p(x)$ is the minimal polynomial for a over F, then

$$F(a) \cong F[x]/\langle p(x) \rangle.$$

Moreover, if $n = \deg(p)$, then every element of $F(a)$ has a unique expression

$$c_0 + c_1 a + c_2 a^2 + \cdots + c_{n-1} a^{n-1}, \quad c_i \in F.$$

Thus $\{1, a, a^2, \ldots, a^{n-1}\}$ spans $F(a)$ over F, and uniqueness gives linear independence. This set is therefore a basis for $F(a)$ as an F-vector space, and

$$[F(a) : F] = \deg(p).$$

Definition 28.9 An extension of the form $F \subseteq F(a)$ is called a **simple extension**.

Example 28.10 We have

$$[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = \deg(x^2 - 2) = 2,$$

since $x^2 - 2$ is irreducible by Eisenstein's criterion. Similarly,

$$[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = \deg(x^3 - 2) = 3,$$

again by Eisenstein's criterion. Also,

$$[\mathbb{C} : \mathbb{R}] = [\mathbb{R}(i) : \mathbb{R}] = \deg(x^2 + 1) = 2.$$

Finally, $[\mathbb{Q}(x) : \mathbb{Q}] = \infty$, since $\{1, x, x^2, x^3, \ldots\}$ is an infinite linearly independent set over $\mathbb{Q}$.

Proposition 28.11 For a field extension $F \subseteq K$, we have $[K : F] = 1$ if and only if $K = F$.

Proof. If $K = F$, then $[K : F] = 1$ because $\{1\}$ is a basis for K over F. Conversely, if $[K : F] = 1$, then $\{1\}$ is a basis for K over F. Thus every element of K is a scalar multiple of 1, so $K = F$. □

Theorem 28.12 Finite field extensions are algebraic.

Proof. Suppose that $[K : F] = n$. If $a \in K$, then the $n + 1$ elements

$$1, a, a^2, \ldots, a^n$$

are linearly dependent over F. Thus there are coefficients $c_0, c_1, \ldots, c_n \in F$, not all zero, such that

$$c_0 + c_1 a + \cdots + c_n a^n = 0.$$

Therefore a is algebraic over F. □

Lecture 29 — Friday, March 18, 2016

The Tower Theorem and Algebraic Extensions

Theorem 28.12 can also be seen directly as follows. If $[K : F] = n < \infty$ and $a \in K$, then the set $\{1, a, a^2, \ldots, a^n\}$ has $n+1$ elements in an n-dimensional vector space over F. Hence it is linearly dependent over F, so a satisfies a nonzero polynomial with coefficients in F.

Example 29.1 Let

$$K = \mathbb{Q}(\sqrt{2}, \sqrt[3]{2}, \sqrt[4]{2}, \sqrt[5]{2}, \ldots).$$

This is an algebraic extension of $\mathbb{Q}$, because the elements algebraic over $\mathbb{Q}$ form a field by Corollary 30.1. However, it is not a finite extension. For each $n \in \mathbb{N}$, the field K contains $\mathbb{Q}(\sqrt[n]{2})$, and $x^n - 2$ is irreducible over $\mathbb{Q}$ by Eisenstein's criterion. Hence

$$[K : \mathbb{Q}] \geq [\mathbb{Q}(\sqrt[n]{2}) : \mathbb{Q}] = n$$

for every $n \in \mathbb{N}$. Therefore $[K : \mathbb{Q}]$ is infinite.

Theorem 29.2 (Tower theorem) Let $F \subseteq E \subseteq K$ be field extensions. If K is finite over E and E is finite over F, then

$$[K : F] = [K : E][E : F].$$

The two stages and their multiplicative degree count are shown in Figure 11.

Proof. Let $X = \{x_1, x_2, \ldots, x_n\}$ be a basis for K over E, and let $Y = \{y_1, y_2, \ldots, y_m\}$ be a basis for E over F. We claim that

$$XY = \{x_i y_j : 1 \leq i \leq n,\ 1 \leq j \leq m\}$$

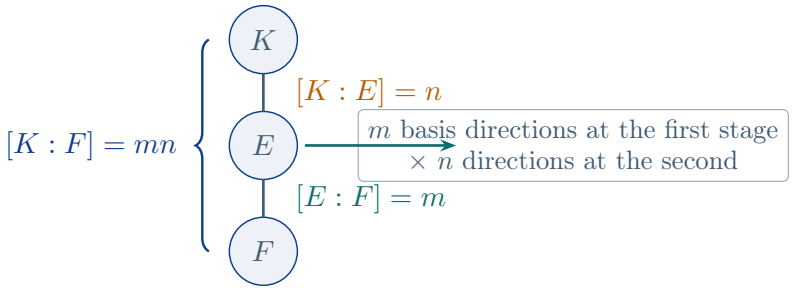


FIGURE 11

is a basis for K over F.

First let $a \in K$. Since X is a basis for K over E, there are $b_i \in E$ such that

$$a = \sum_{i=1}^{n} b_i x_i.$$

Since Y is a basis for E over F, for each i there are $c_{ij} \in F$ such that

$$b_i = \sum_{j=1}^{m} c_{ij} y_j.$$

Combining these expressions gives

$$a = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} c_{ij} y_j \right) x_i = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} x_i y_j.$$

Thus XY spans K over F.

Now suppose that

$$0 = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} x_i y_j = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} c_{ij} y_j \right) x_i,$$

where each $c_{ij} \in F$. Since $\sum_{j=1}^{m} c_{ij} y_j \in E$ and X is linearly independent over E, we have

$$\sum_{j=1}^{m} c_{ij} y_j = 0$$

for each i. Since Y is linearly independent over F, it follows that $c_{ij} = 0$ for all i and j. Thus XY

is linearly independent over F. Therefore XY is a basis for K over F, and

$$[K : F] = |XY| = nm = [K : E][E : F].$$

□

Example 29.3 We compute $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}]$ and find a basis. By Theorem 29.2,

$$[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{2})(\sqrt{3}) : \mathbb{Q}(\sqrt{2})]\,[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}].$$

Since $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$, it remains to compute the first factor. The element $\sqrt{3}$ is a root of $x^2 - 3 \in \mathbb{Q}(\sqrt{2})[x]$. If $p(x)$ is the minimal polynomial for $\sqrt{3}$ over $\mathbb{Q}(\sqrt{2})$, then $p(x) \mid x^2 - 3$, so

$$[\mathbb{Q}(\sqrt{2})(\sqrt{3}) : \mathbb{Q}(\sqrt{2})] = \deg(p) \leq 2.$$

We claim that this degree is not 1. If it were 1, then $\sqrt{3} \in \mathbb{Q}(\sqrt{2})$, so

$$\sqrt{3} = a + b\sqrt{2}$$

for some $a, b \in \mathbb{Q}$. Squaring gives

$$3 = a^2 + 2b^2 + 2ab\sqrt{2}.$$

If $a = 0$, then $3 = 2b^2$, impossible for $b \in \mathbb{Q}$. If $b = 0$, then $3 = a^2$, impossible for $a \in \mathbb{Q}$. If $a \neq 0$ and $b \neq 0$, then

$$\sqrt{2} = \frac{3 - a^2 - 2b^2}{2ab} \in \mathbb{Q},$$

again a contradiction. Therefore

$$[\mathbb{Q}(\sqrt{2})(\sqrt{3}) : \mathbb{Q}(\sqrt{2})] = 2,$$

and hence $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = 4$.

A basis for $\mathbb{Q}(\sqrt{2})$ over $\mathbb{Q}$ is $\{1, \sqrt{2}\}$, and a basis for $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ over $\mathbb{Q}(\sqrt{2})$ is $\{1, \sqrt{3}\}$. The tower theorem construction therefore gives the basis

$$\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$$

for $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ over $\mathbb{Q}$.

Example 29.4 Let

$$K = \mathbb{Q}(\sqrt[3]{2}, \sqrt[4]{3}) \qquad \text{and} \qquad n = [K : \mathbb{Q}].$$

Both generators are algebraic over $\mathbb{Q}$, so adjoining them successively shows that $K/\mathbb{Q}$ is finite. By Theorem 29.2,

$$n = [K : \mathbb{Q}(\sqrt[3]{2})]\,[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3[K : \mathbb{Q}(\sqrt[3]{2})],$$

since $x^3 - 2$ is irreducible by Eisenstein's criterion. Similarly,

$$n = [K : \mathbb{Q}(\sqrt[4]{3})]\,[\mathbb{Q}(\sqrt[4]{3}) : \mathbb{Q}] = 4[K : \mathbb{Q}(\sqrt[4]{3})],$$

since $x^4 - 3$ is irreducible by Eisenstein's criterion. Thus $3 \mid n$ and $4 \mid n$, so $12 \mid n$ and $n \geq 12$.

On the other hand, $\sqrt[3]{2}$ is a root of $x^3 - 2 \in \mathbb{Q}(\sqrt[4]{3})[x]$, so $[K : \mathbb{Q}(\sqrt[4]{3})] \leq 3$. Therefore

$$n = [K : \mathbb{Q}(\sqrt[4]{3})]\,[\mathbb{Q}(\sqrt[4]{3}) : \mathbb{Q}] \leq 3 \cdot 4 = 12.$$

Hence $[K : \mathbb{Q}] = 12$.

Corollary 29.5 If $F \subseteq E \subseteq K$, K is finite over E, and E is finite over F, then K is finite over F.

Proof. This is immediate from Theorem 29.2 because $[K : F] = [K : E][E : F]$ is a product of two finite numbers. □

Corollary 29.6 If a is transcendental over F, then $[F(a) : F] = \infty$.

Proof. If $[F(a) : F]$ were finite, then $F(a)$ would be algebraic over F by Theorem 28.12. This would imply that a is algebraic over F, a contradiction. □

Theorem 29.7 If $F \subseteq E \subseteq K$, K is algebraic over E, and E is algebraic over F, then K is algebraic over F.

Proof. Let $a \in K$. Since K is algebraic over E, there is a nonzero polynomial $p(x) \in E[x]$ such that $p(a) = 0$. Write

$$p(x) = b_0 + b_1 x + \cdots + b_n x^n, \quad b_i \in E.$$

Since E is algebraic over F, each coefficient b_i is algebraic over F. Define

$$F_0 = F(b_0), \qquad F_1 = F_0(b_1), \qquad F_2 = F_1(b_2), \qquad \ldots, \qquad \text{and} \qquad F_n = F_{n-1}(b_n).$$

Each extension F_i over F_{i-1} is finite, because b_i is algebraic over F_{i-1}. By repeated use of Theorem 29.2, the extension F_n over F is finite.

Now $p(x) \in F_n[x]$ and $p(a) = 0$, so a is algebraic over F_n. Therefore $F_n(a)$ is finite over F_n. Applying Theorem 29.2 once more, $F_n(a)$ is finite over F. By Theorem 28.12, the extension $F_n(a)$ over F is algebraic. Since $a \in F_n(a)$, it follows that a is algebraic over F. □

Lecture 30 — Monday, March 21, 2016

Recall three facts about field extensions. By Theorem 28.12, finite extensions are algebraic. If a is algebraic over F, then the simple extension $F(a)$ is finite over F. Finally, by Theorem 29.7, if $F \subseteq E \subseteq K$, K is algebraic over E, and E is algebraic over F, then K is algebraic over F.

Corollary 30.1 Let $F \subseteq K$ be a field extension. Then

$$L = \{a \in K : a \text{ is algebraic over } F\}$$

is a subfield of K.

Proof. Suppose that $a, b \in L$ with $b \neq 0$. It is enough to show that $a - b$, ab, and b^{-1} are algebraic over F.

Since a is algebraic over F, the extension $F(a)$ is finite over F. Since b is algebraic over F, it is also algebraic over $F(a)$, so $F(a, b) = F(a)(b)$ is finite over $F(a)$. By Theorem 29.2, the extension $F(a, b)$ is finite over F. Hence $F(a, b)$ is algebraic over F by Theorem 28.12. Since $a - b$, ab, and b^{-1} all lie in $F(a, b)$, they are algebraic over F. Therefore L is a subfield of K. □

5. Finite Fields and Primitive Elements

Finite Fields

Finite fields were introduced by Évariste Galois in 1830, while he was alive. His motivation was to understand the solvability of polynomial equations, and the quintic in particular. However, finite fields also have important applications in areas such as cryptography. Our goal is to classify all finite fields.

Every finite field has order equal to a prime power. Indeed, if K is a finite field, then its characteristic is a prime p, so it contains a copy of $\mathbb{Z}_p$. As a finite-dimensional vector space over $\mathbb{Z}_p$, it has order p^n for some $n \in \mathbb{N}$.

Definition 30.2 Let

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x],$$

where F is a field. The **formal derivative** of $f(x)$ is

$$f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1.$$

The usual product rule holds for formal derivatives.

Lemma 30.3 A polynomial $f(x) \in F[x]$ has a multiple root in some extension E of F if and only if $f(x)$ and $f'(x)$ have a common factor of positive degree in $F[x]$.

Proof. *Forward implication.* Suppose that $f(x)$ has a multiple root in some extension E of F. If a is that multiple root, then

$$f(x) = (x-a)^2 g(x)$$

for some $g(x) \in E[x]$. By the product rule,

$$f'(x) = (x-a)^2 g'(x) + 2(x-a)g(x),$$

so $(x-a) \mid f'(x)$ in $E[x]$.

Assume, for a contradiction, that $f(x)$ and $f'(x)$ have no common factor of positive degree in $F[x]$.

Since $F[x]$ is a Euclidean domain, there are $h(x), k(x) \in F[x]$ such that

$$f(x)h(x) + f'(x)k(x) = 1.$$

Viewing this identity in $E[x]$, the polynomial $(x-a)$ divides the left-hand side, and hence divides 1, a contradiction. Thus $f(x)$ and $f'(x)$ have a common factor of positive degree in $F[x]$.

Reverse implication. Suppose that $f(x)$ and $f'(x)$ have a common factor of positive degree in $F[x]$. Let a be a root of this common factor in some extension E of F. Then, in $E[x]$,

$$f(x) = (x-a)g(x)$$

for some $g(x) \in E[x]$, and $(x-a) \mid f'(x)$. By the product rule,

$$f'(x) = (x-a)g'(x) + g(x).$$

Since $(x-a) \mid f'(x)$, it follows that $(x-a) \mid g(x)$. Hence $g(x) = (x-a)h(x)$ for some $h(x) \in E[x]$, so

$$f(x) = (x-a)^2 h(x).$$

Therefore a is a multiple root of $f(x)$. □

Theorem 30.4 (Classification theorem for finite fields) For each prime p and each $n \in \mathbb{N}$, there is a unique finite field of size p^n, up to isomorphism.

Proof. Let E be the splitting field of

$$f(x) = x^{p^n} - x$$

over $\mathbb{Z}_p$. Since

$$f'(x) = p^n x^{p^n - 1} - 1 = -1$$

in $\mathbb{Z}_p[x]$, Lemma 30.3 shows that $f(x)$ has no multiple roots. Thus $f(x)$ has exactly p^n distinct roots in E. Let

$$K = \{a \in E : f(a) = 0\}.$$

Then $|K| = p^n$.

We show that K is a subfield of E. If $a, b \in K$, then in characteristic p we have

$$(a+b)^{p^n} = a^{p^n} + b^{p^n} = a + b,$$

so $a+b \in K$. Also,

$$(-a)^{p^n} = -a$$

and

$$(ab)^{p^n} = a^{p^n} b^{p^n} = ab,$$

so $-a \in K$ and $ab \in K$. If $a \in K$ and $a \neq 0$, then

$$(a^{-1})^{p^n} = (a^{p^n})^{-1} = a^{-1},$$

so $a^{-1} \in K$. Thus K is a subfield of E. Since $f(x)$ splits over K, the minimality of the splitting field gives $K = E$. Therefore there is a finite field of order p^n.

For uniqueness, keep E as above, and let L be any finite field of order p^n. The multiplicative group $L \setminus \{0\}$ has order $p^n - 1$. By Lagrange's theorem,

$$a^{p^n - 1} = 1$$

for every nonzero $a \in L$, and hence $a^{p^n} = a$ for every $a \in L$. Thus every element of L is a root of $f(x) = x^{p^n} - x$. Since $|L| = \deg(f)$, the polynomial $f(x)$ splits over L, and L is a splitting field for $f(x)$ over $\mathbb{Z}_p$. By Corollary 27.4, we have $L \cong E$. Hence the finite field of order p^n is unique up to isomorphism. □

Lecture 31 — Wednesday, March 23, 2016

Subfields of Finite Fields

By Theorem 30.4, for every prime p and every $n \in \mathbb{N}$ there is a unique field of order p^n up to isomorphism. It is the splitting field over $\mathbb{Z}_p$ of $x^{p^n} - x$, and its elements are precisely the roots of this polynomial.

Definition 31.1 The unique field of order p^n is denoted by $GF(p^n)$ and is called the **Galois field of order p^n**, in honor of Galois.

Proposition 31.2 As an additive group,

$$GF(p^n) \cong \underbrace{\mathbb{Z}_p \oplus \mathbb{Z}_p \oplus \cdots \oplus \mathbb{Z}_p}_{n \text{ factors}}.$$

Proof. The field $GF(p^n)$ contains its prime subfield $GF(p) \cong \mathbb{Z}_p$. Hence it is a vector space over $GF(p)$. If its dimension over $GF(p)$ is d, then it has p^d elements. Since $|GF(p^n)| = p^n$, we have $d = n$.

Choosing a basis identifies the additive group of $GF(p^n)$ with the additive group of the n-dimensional vector space over $\mathbb{Z}_p$, which is $\mathbb{Z}_p^n$. □

Corollary 31.3 We have $[GF(p^n) : GF(p)] = n$.

Proof. This is the vector space dimension computed in Proposition 31.2. □

Theorem 31.4 For every divisor m of n, there is a unique subfield of $GF(p^n)$ of order p^m. Moreover, these are the only subfields of $GF(p^n)$.

Proof. Suppose first that $m \mid n$. Then

$$p^n - 1 = (p^m - 1)(p^{n-m} + p^{n-2m} + \cdots + p^m + 1) = (p^m - 1)t,$$

for some positive integer t. Let

$$K = \{x \in GF(p^n) : x^{p^m} - x = 0\}.$$

As in the proof of Theorem 30.4, the set K is a subfield of $GF(p^n)$.

We next recall why the multiplicative group of a finite field is cyclic. Let

$$G = GF(p^n) \setminus \{0\}.$$

By the fundamental theorem of finite abelian groups,

$$G \cong \mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \cdots \oplus \mathbb{Z}_{n_k}.$$

Suppose that there were an integer $d > 1$ such that $d \mid n_i$ and $d \mid n_j$ for some $i \neq j$. Then G would have two distinct subgroups H_1 and H_2 of order d. Every element of each subgroup is a root of $x^d - 1$, by Lagrange's theorem. Since $|H_1 \cup H_2| > d$, this would give more than d roots of $x^d - 1$ in the field $GF(p^n)$, which is impossible. Thus the integers $n_1, \ldots, n_k$ are pairwise coprime, and so G is cyclic.

Choose a generator a of G. Since a has order $p^n - 1$ and t divides $p^n - 1$, the element a^t has order

$$\frac{p^n - 1}{t} = p^m - 1.$$

Therefore the subgroup $\langle a^t \rangle$ has $p^m - 1$ elements, and every element $u \in \langle a^t \rangle$ satisfies

$$u^{p^m - 1} = 1, \qquad \text{so} \qquad u^{p^m} = u.$$

Thus

$$\langle a^t \rangle \cup \{0\} \subseteq K.$$

Hence $|K| \geq p^m$. Since K is the set of roots of the polynomial $x^{p^m} - x$, it has at most p^m elements. Therefore $|K| = p^m$, and K is a subfield of $GF(p^n)$ of order p^m.

For uniqueness, let H be any subfield of $GF(p^n)$ of order p^m. If $u \in H$ is nonzero, then $u^{p^m - 1} = 1$ by Lagrange's theorem, so $u^{p^m} = u$. Also $0^{p^m} = 0$. Thus every element of H is a root of $x^{p^m} - x$, so $H \subseteq K$. Since $|H| = |K| = p^m$, we have $H = K$.

Finally, suppose that H is any subfield of $GF(p^n)$. Then H contains $GF(p)$, so $|H| = p^m$ for some positive integer m. By Theorem 29.2,

$$n = [GF(p^n) : GF(p)] = [GF(p^n) : H]\,[H : GF(p)] = [GF(p^n) : H]\,m.$$

Thus $m \mid n$. □

Example 31.5 The subfield lattice of $GF(2^{24})$ is determined by the divisors

$$1,\ 2,\ 3,\ 4,\ 6,\ 8,\ 12,\ 24$$

of 24.

The corresponding subfield lattice is shown in Figure 12.

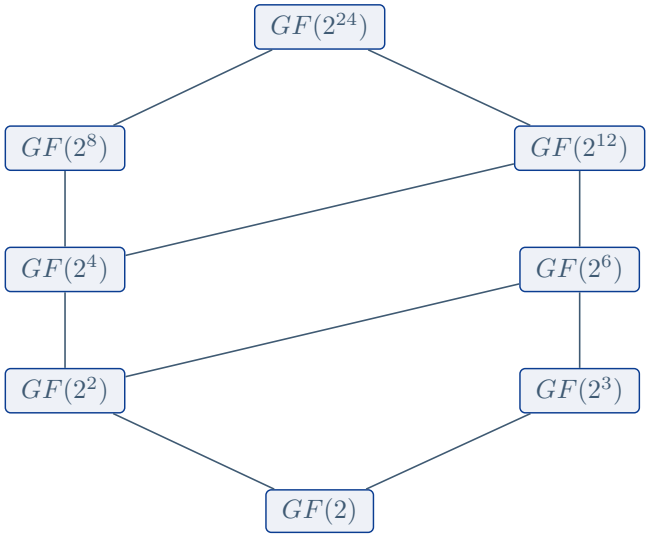


Figure 12

Lecture 32 — Monday, March 28, 2016

Primitive Elements

We continue to use the relationship between finite and algebraic extensions: finite extensions are algebraic, and simple algebraic extensions are finite.

Definition 32.1 Let $F \subseteq K$ be a field extension. An element $\alpha \in K$ is called a **primitive element** for the extension if $K = F(\alpha)$.

Theorem 32.2 (Primitive element theorem) Let F be a field of characteristic 0, and suppose that a and b are algebraic over F. Then there is an element $c \in F(a, b)$ such that

$$F(a, b) = F(c).$$

Proof. Let $p(x)$ and $q(x)$ be the minimal polynomials of a and b over F, respectively. In a

common splitting field, write the distinct roots of $p(x)$ as

$$a = a_1, a_2, \ldots, a_n$$

and the distinct roots of $q(x)$ as

$$b = b_1, b_2, \ldots, b_m.$$

Since F has characteristic 0, irreducible polynomials over F have no repeated roots.

If $m = 1$, then $q(x)$ has only one root and no repeated roots, so $q(x)$ is linear. Hence $b \in F$, and $F(a, b) = F(a)$.

Now suppose that $m > 1$. Since F has characteristic 0, it is infinite. Therefore we may choose $d \in F$ such that

$$d \neq \frac{a_i - a}{b - b_j}$$

for all $1 \leq i \leq n$ and $2 \leq j \leq m$. Equivalently,

$$a_i \neq a + d(b - b_j)$$

for all $1 \leq i \leq n$ and $2 \leq j \leq m$.

Set

$$c = a + db \in F(a, b).$$

We claim that $F(c) = F(a, b)$. The inclusion $F(c) \subseteq F(a, b)$ is immediate. To prove the reverse inclusion, define

$$r(x) = p(c - dx) \in F(c)[x].$$

Then

$$r(b) = p(c - db) = p(a) = 0.$$

Let $s(x)$ be the minimal polynomial of b over $F(c)$. Since $r(b) = 0$ and $q(b) = 0$, Proposition 28.5 gives

$$s(x) \mid r(x) \qquad \text{and} \qquad s(x) \mid q(x)$$

in $F(c)[x]$.

The roots of $s(x)$ in the chosen splitting field must therefore be common roots of $r(x)$ and $q(x)$. For $j > 1$,

$$r(b_j) = p(c - db_j) = p(a + d(b - b_j)).$$

By the choice of d, the element $a + d(b - b_j)$ is not one of the roots a_i of $p(x)$. Hence $r(b_j) \neq 0$ for $j > 1$. Thus the only common root of $r(x)$ and $q(x)$ is $b = b_1$.

Since $s(x)$ is irreducible over a field of characteristic 0, it has no repeated roots. Therefore $s(x)$ is linear, so $s(x) = x - b$. Hence $b \in F(c)$. Since $a = c - db$, we also have $a \in F(c)$. Thus $F(a, b) \subseteq F(c)$, and the proof is complete. □

Corollary 32.3 Suppose that F has characteristic 0 and that $a_1, \ldots, a_n$ are algebraic over F. Then there is an element $c \in F(a_1, \ldots, a_n)$ such that

$$F(a_1, \ldots, a_n) = F(c).$$

Proof. We argue by induction on n. The case $n = 1$ is immediate. Suppose that the result holds for $n - 1$ algebraic elements. Then there is an element $d \in F(a_1, \ldots, a_{n-1})$ such that

$$F(a_1, \ldots, a_{n-1}) = F(d).$$

Since d and a_n are algebraic over F, Theorem 32.2 gives an element $c \in F(d, a_n)$ such that

$$F(d, a_n) = F(c).$$

Therefore

$$F(a_1, \ldots, a_n) = F(d, a_n) = F(c).$$

□

Remark 32.4 The proof of Theorem 32.2 used the characteristic 0 hypothesis in two places. First, it used the fact that a field of characteristic 0 is infinite. Second, it used the fact that irreducible polynomials over a field of characteristic 0 have no repeated roots in any extension field.

More generally, the primitive element theorem holds over any field for which no irreducible polynomial has a repeated root in any extension. Such fields are called **perfect fields**.

The primitive element theorem also holds for finite fields. Suppose that F is finite and that

$$K = F(a, b),$$

where a and b are algebraic over F. Then K is finite by Theorem 29.2. As shown in the proof of Theorem 31.4, the multiplicative group $K \setminus \{0\}$ is cyclic. If

$$K \setminus \{0\} = \langle \alpha \rangle,$$

then $F(\alpha)$ contains every nonzero element of K, as well as 0. Hence $F(\alpha) = K$.

Appendix: Lecture and Tutorial Index